(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 January 2004 (08.01.2004)

PCT

(10) International Publication Number
WO 2004/003159 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number:
PCT/US2003/020410

(22) International Filing Date: 27 June 2003 (27.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/391,878 27 June 2002 (27.06.2002) US

(71) Applicants *(for all designated States except US)*: **XENON GENETICS, INC.** [CA/CA]; 3650 Gilmore Way, Burnaby, British Columbia V5G 4W8 (CA). **WARNER-LAMBERT COMPANY, LLC** [US/US]; 201 Tabor Road, Morris Plains, NJ 07950 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BROWNLIE, Alison, J.** [CA/CA]; 1823 Napier, Vancouver, British Columbia V5L 2N4 (CA). **TAFURI,Sherrie, Rae** [US/US]; 4773 Aberdeen Drive, Ann Arbor, MI 48103 (US). **BRINKMAN, Ryan, R.** [CA/CA]; 627 West 7th Street, Vancouver, British Columbia V5Z 1B6 (CA). **CHAGNOVICH, Daniel** [US/US]; 2720 Aspen Ridge Drive, Ann Arbor, MI 48103 (US). **CHATTERJEE, Aurobindo** [IN/IN]; 712 Mt. Pleasant Avenue, Ann Arbor, MI 48103 (US). **DONALDSON, Gary, C.** [CA/CA]; 17763 - 24th Avenue, Surrey, British Columbia V3S 9V2 (CA). **DUBE, Marie-Pierre** [CA/CA]; 4831 Hutchison Street, Montreal, Quebec H2V 4A4 (CA). **GOLDBERG, Yigal, P.** [CA/CA]; 62 West 18th Avenue, Vancouver, British Columbia V5Y 2A4 (CA). **JERVA, Leonard, F.** [US/US]; 2910 Green Valley Drive, Ann Arbor, MI 48103 (US). **LAFRENIER, Ronald, G.** [CA/—]; 1264 Osborne Avenue, Verdun, Quebec H4H 1X5 (CA). **LUDWIG, Erwin** [CA/CA]; 4410 Lateener Lane, Blaine, WA 98230 (US). **SAMUELS, Mark, E.** [US/US]; 2702 Manitoba Street, Vancouver, British Columbia V5Y 3Y9 (CA). **WU, Chenyan** [CN/CN]; 3190 Otter Creek Court, Ann Arbor, MI 48105 (US). **HAYDEN, Michael, R.** [US/CA]; 4484 West 7th Avenue, Vancouver, British Columbia V6R 1W9 (CA).

(74) Agents: **GRANT, Alan, J.** et al.; Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07068 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY; TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/003159 A2

(54) Title: NOVEL THERAPEUTIC TARGET FOR TREATING VASCULAR DISEASES, DYSLIPIDEMIAS AND RELATED DISORDERS

(57) Abstract: The invention features compositions and methods for modulating HDL levels and/or treating a vascular disease, dyslipidemia or a related disorder in a patient in need thereof. The invention also includes methods for identifying therapeutic agents for such treating disorders based on the identification of novel therapeutic targets herein. The invention also includes diagnostic and pharmacogenomic compositions and methods which employ the therapeutic targets named herein.

NOVEL THERAPEUTIC TARGET FOR TREATING VASCULAR DISEASES, DYSLIPIDEMIAS AND RELATED DISORDERS

This application claims priority of U.S. Provisional Application Serial No. 60/391,878, filed 27 June 2002, the disclosure of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to the discovery of gene95, a novel therapeutic target for disease diagnosis and treatment, especially diagnosis and treatment of cardiovascular diseases, such as dyslipidemia and low HDL diseases, including hypoalphalipoproteinemia, as well as therapeutic agents and methods of screening and identifying therapeutic agents useful in preventing, treating or otherwise ameliorating such disorders.

BACKGROUND OF THE INVENTION

Epidemiological studies have consistently demonstrated that plasma high density lipoprotein cholesterol (HDL-C) concentration is inversely related to the incidence of vascular disease, particularly cardiovascular disease

(CVD) and coronary artery disease (CAD). HDL levels are a strongly graded and independent cardiovascular risk factor. Protective effects of an elevated HDL-C persist until 80 years of age. A low HDL-C is associated with an increased CAD risk even with normal (<5.2 mmol/l) total plasma cholesterol levels. Even in the face of other dyslipidemias or secondary factors, HDL-C levels are important predictors of CAD. Low HDL cholesterol (in severe cases called hypoalphalipoproteinemia), is also implicated in cerebrovascular disease, coronary restenosis, and peripheral vascular disease.

Pharmacological intervention of low HDL-C levels has so far proven unsatisfactory. Currently, there are no FDA approved therapeutic agents which modulate HDL levels in a direct, significant fashion. Current research strategies aimed towards modulating HDL involve increasing production of ApoA1, promoting the rate of reverse cholesterol transport (RCT) and decreasing catabolism of HDL. For example, increasing ApoA1 levels may be possible via small molecular upregulation of transcription or by infusion of ApoA1 protein. Alternatively, it may be possible to upregulate RCT by developing ABCA1 agonists and increasing cholesterol efflux from peripheral tissues (see PCT publication WO 00/55318, incorporated herein by reference). Catabolism of HDL is regulated by a number of enzymes, including Cholesteryl Ester Transfer Protein (CETP), which may also be a suitable therapeutic target for HDL regulation.

Identification of additional genes and corresponding proteins which act to modulate HDL levels in humans would allow rational choice of which therapeutic approach to pursue in terms of drug development. The present invention relates to the discovery of a new target for therapeutic intervention in the treatment of vascular diseases and dyslipidemias and use of said target to screen for and identify therapeutic agents useful in alleviating low HDL levels and the symptoms of attendant disorders.

BRIEF SUMMARY OF THE INVENTION

The present invention provides the complete *gene95* cDNA nucleotide sequence (SEQ ID NO: 3) and gene95 amino acid sequence (SEQ ID NO: 4) from humans. The invention further identifies the role and function of gene95 as a key regulator of HDL cholesterol levels by identification of mutant forms of the gene which are associated with inherited low HDL disorders in humans.

In one aspect, the present invention provides the purified nucleic acid sequence of *gene95* and fragments thereof, including purified and semi-purified forms thereof, as well as recombinant cell lines, viruses or cell extracts containing *gene95* nucleotide constructs.

In another aspect, the present invention provides the purified gene95 protein and fragments thereof, including purified and semi-purified forms thereof, as well as cells or cell extracts containing gene95 biological activity, and anti-gene95 antibodies, including methods of making such polypeptides.

In one aspect, the present invention provides a method for determining whether a candidate compound modulates HDL levels and/or treating a vascular disease or dyslipidemia. The method comprises (a) determining a biological activity of the gene95 gene or protein; (b) contacting a source of this biological activity with a candidate compound under conditions promoting this biological activity; and (c) determining a change in the biological activity of said gene or protein as a result of said contacting, wherein the change in biological activity identifies the candidate compound, or ~~its~~ an analog of the candidate compound, as a compound that modulates said biological activity and is useful in regulating HDL levels and/or treating a vascular disease, dyslipidemia or a related disorder. Other types of assays for Gene95 activity are also specifically contemplated by the present invention.

In another aspect, the invention provides a method for computationally identifying a compound for modulating HDL levels and/or treating a vascular disease or dyslipidemia. The method involves (a) determining the active site of the gene95 protein selected (for example, through X-Ray crystallography or other techniques); and (b) through computational modeling, identifying a compound which interacts with the active site, thereby identifying a compound, or an analog thereof, as a compound which is useful for modulating HDL levels and/or treating a vascular disease, dyslipidemia or a related disorder.

In another aspect, the invention provides a method for modulating HDL levels and/or treating a vascular disease, dyslipidemia or a related disorder, comprising administering to a person in need thereof, a compound which modulates the activity of gene 95 gene or protein.

In a further aspect, the invention provides tools for diagnosis of vascular disease and dyslipidemias and/or pharmacogenomics of therapeutic agents for treating vascular disease and dyslipidemias, comprising identification of mutations in the gene95 gene or protein, for example the mutations identified in SEQ ID Nos. 6 – 9 or in Table 1.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the pedigree of family BC-11M.

Figure 2 shows the pedigree of family NL-003.

Figure 3 shows the pedigree of family NL-120.

Figure [illegible] illustrates the expected 10 exon construct of gene95 cDNA. Incomplete constructs assembled from known fragments of gene95 are also illustrated, including other possible transcripts employing E6a or E9a.

Figures 5A-5G provide the sequence of exons and introns and the amino acid translation of gene95. Legend: Lower case letters = intron regions not forming part of mature mRNA or cDNA. "..." = intron regions available from public database; not included herein. Capital Letters = Nucleotide base or Amino Acid (as indicated by context; with standard designations used). "*" = a stop codon. +1, +2, +3: Frameshift for translation of nucleic acid. Short open reading frames before and after the correct amino acid sequence are indicated for completeness, but are not expected to form part of the gene95 protein. Here, exon E1 is part of SEQ ID NO: 18, exon E2 is part of SEQ ID NO: 19, exon E3 is part of SEQ ID NO: 20, exon E4 is part of SEQ ID NO: 21, exon E5 is part of SEQ ID NO: 22, exon E6 is part of SEQ ID NO: 23, exon E7 is part of SEQ ID NO: 24, exon E8 is part of SEQ ID NO: 25, exon E9 is part of SEQ ID NO: 26, exon E10 is part of SEQ ID NO: 27, exon E6a is part of SEQ ID NO: 28, and exon E9a is part of SEQ ID NO: 29. Here, exon 6a = Alternative Exon 6 ("Exon 6 alt") (truncates protein) while exon 9a = Alternative exon 9 ("Exon 9 alt") (longer, but truncates). In addition, ctttga (SEQ ID NO: 30) shows the start of a new intronic sequence.

Figure 6 illustrates the alignment of human (SEQ ID NO: 4) and mouse (SEQ ID NO: 5) amino acid sequences of gene95.

Figure 7 shows the alignments of the amino acid sequence of mutation 1 (SEQ ID NO: 7) and mutation 2 (SEQ ID NO: 16) with the wild–type amino acid sequence (SEQ ID NO: 4) of human gene95.

Figure 8 illustrates the predicted exon assembly of gene95, with locations of mutations and polymorphisms identified according to the invention.

Figure 9 shows a Northern blot with a Gene 95 probe illustrating the results of Example 3.

DEFINITIONS

As used herein, the following terms have their indicated meanings unless expressly stated otherwise.

As used herein, the term "correspond" in the sense of "a gene corresponding to gene95", means that the gene has the indicated nucleotide sequence or that it encodes substantially the same RNA as would be encoded by the indicated sequence, the term "substantially" meaning at least about 50% identical over a sequence encoding a protein with gene95 biological activity; more preferably at least about 60%, 70%, 80%, 90% or higher, and as defined elsewhere herein, including splice variants thereof.

"Corresponding to gene95" also includes homologs or orthologs of gene95 from another organism, and sequence variants thereof having at least about 50% identity, at either the nucleotide or amino acid level, and more preferably at least about 60%, 70%, 80%, 90% or higher homology. Preferably such gene is from a eukaryote, more preferably a mammal, more preferably a rodent or human. The human genes and proteins listed in this specification are representative of a wide variety of homologs, orthologs and analogs that may be obtained from other species. These homologs, orthologs and/or analogs, and reasonable variants thereof, can be used in the compositions and methods of this invention. For example, a screening assay which uses the mouse homolog of a listed human gene may also be used in a screening assay to identify potential therapeutic agents for treating humans.

As such, the present invention relates to the use of these isoforms, homologs from other species, and analogous sequences which, when translated, retain the biological activity of the isoform, as determined on a region by region basis. Such analogous sequences share preferably greater than 50% identity with SEQ ID NO: 3 or SEQ ID NO: 4, more preferably 60%, 70%, 80% and most preferably at least about 90% identity with the isoforms set out herein. All such sequences may still be useful in practicing the methods of the invention provided that they retain sufficient therapeutic target biological activity so as to determine the effects of test compounds in regulating such activity. Such understanding is available to those skilled in the art based on the teachings disclosed herein.

Further in accordance with the present invention, the term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

$$\text{Percent Identity} = 100\,[1-(C/R)]$$

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence for which the percent identity as calculated above is about equal to or greater than a specified minimum Percent Identity then the Compared Sequence has the specified minimum percent identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

As used herein, the terms "portion," "segment," and "fragment," when used in relation to polypeptides, refer to a continuous sequence of nucleotide residues, sequence forms a subset of a larger sequence. Such terms include the products produced by treatment of said polynucleotides with any of the common endonucleases, or any stretch of polynucleotides that could be synthetically synthesized. These may include exonic and intronic sequences of the corresponding genes.

"Therapeutic target biological activity" (or just "biological activity") as used herein is a very broad term that relates to all the directly or indirectly measurable and identifiable biological activities of the therapeutic target gene and protein, including but not limited to those examples set out in (a) – (d) below:

(a) Relating to the purified therapeutic target protein, therapeutic target biological activity includes, but is not limited to, all those biological processes, interactions, or binding of ligands, proteins, membrane components or other compounds (such as small organic compounds), binding behavior, binding-activity relationships, pKa, pD, enzyme kinetics, stability, and functional assessments of the protein.

(b) Relating to therapeutic target biological activity in cell fractions, reconstituted cell fractions or whole cells, these activities include, but are not limited to the ligand or antibody binding behavior and all measurable

consequences of this effect, such as measurement of any signaling cascade, efflux, influx or accumulation, stability or degradation of cellular components, membrane composition and behavior, cell growth, development or behavior and other direct or indirect effects of therapeutic target activity.

(c) Relating to therapeutic target genes and transcription, therapeutic target biological activity includes the rate, scale or scope of transcription of DNA to generate therapeutic target mRNA or its alternate transcripts; the effect of regulatory proteins on such transcription, the effect of modulators of such regulatory proteins on such transcription; plus the stability and behavior of mRNA transcripts, post-transcription processing, mRNA amounts and turnover, and all measurements of translation of the mRNA into polypeptide sequences.

(d) Relating to therapeutic target biological activity in organisms, this includes but is not limited to biological activities which are identified by their absence or deficiency in disease processes or disorders caused by aberrant therapeutic target biological activity in those organisms. Broadly speaking, therapeutic target biological activity can be determined by all these and other means for analyzing biological properties of proteins and genes that are known in the art.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, a gene and protein associated with maintenance of HDL levels in humans has been determined by identification of mutant forms of the gene in families having low HDL levels using the positional cloning strategy set out in Example 1.

The pedigree of family BC-11M is shown in Figure 1. The phenotype in the family used to define the linkage segregates a low HDL (hypoalphalipoproteinemia) trait in an autosomal dominant fashion with high penetrance. Other lipid parameters are normal in this family (including triglyceride (TG) and low density lipoprotein (LDL)) and there is an absence of other confounders (such as obesity and diabetes). There is strong evidence for cardiovascular disease and coronary artery disease in particular in this family as multiple members have had cardiac events (myocardial infarct (MI), stroke and angina) and have been treated for symptoms of this disease (such as with coronary artery bypass graft (CABG)). One member of this family has Behcet syndrome (related to lupus and auto-immune disorders; another member has celiac disease (gluten intolerance) and others have thyroid problems.

Members of this family had the wild type sequence for coding regions of ABCA1, thus distinguishing them from the previously identified low-HDL causing mutations in ABCA1. In addition, the ABCA1 gene is excluded based on the genetics of this family, and based on normal cholesterol efflux measurements in cultured fibroblasts from affected individuals in this family.

The pedigree of family NL-003 is shown in Figure 2. The proband (II:35) has an HDL level below the fifth percentile (age and sex corrected). There are a few cases of overweight/obesity (BMI>25>30) but this condition does not fully co-segregate with the HDL condition and so may not be related.

The pedigree of family NL-120 is shown in Figure 3. Again the proband (II:01) has an HDL level below the fifth percentile (age and sex corrected). In both NL-003 and NL-120 multiple individuals with low HDL are observed. In family NL-003 individuals with low HDL and cardiovascular disease and cardiac events (MI) are observed.

Thus, the genes identified according to the invention are located on chromosome 9, at or near the 9q22 band. Markers for this locus have been identified and are as follows: 9q22ca7 (SEQ ID NO: 1) and 9q22ca17 (SEQ ID NO: 2). Other markers are available in the public domain.

Description of Gene95 and its mutations

Gene95 is alternately known and described herein as P1G95, gene95, G95 or as HDL Regulatory Protein. Aspects of Gene95 are described at a variety of different sources in public databases.

Goldenpath Gene ID: ENSG00000136925, including transcript ID: ENST00000259452. (GoldenPath assembly of the Human Genome Project Working Draft and found at www.genome.ucsc.edu or at the Ensembl web-site www.ensembl.org, version effective August 2001).

GenBank Accession Numbers: AK056453, XM_088573 (Available at www.ncbi.nlm.nih.gov/entrez/query.fcqi). Those skilled in the art will recognize that these references do not provide the correct full length nucleotide coding sequences for gene95, nor do they provide the correct amino acid sequence of gene95 when translated.

The full length cDNA sequence for gene95 (P1G95) is set out in SEQ ID NO: 3 comprising a 4333 bp transcript including 5'UTR and 3' tail. The exon/intron structure of gene95 is set out in Figure 4. The gene comprises 10 exons with possible alternate transcripts for exon 6 ("exon 6 alt" or E6a) and exon 9 ("exon 9 alt" or E9a). Figure 4 illustrates the expected 10 exon construct and other possible transcripts employing E6a or E9a.

The gene95 protein comprises a 516 amino acid protein (SEQ ID NO: 4). Use of an alternate methionine start site may generate a sequence of 489

amino acid analysis of the protein sequence of gene95 suggests it has the characteristics of a soluble protein which is involved in maintaining serum lipid homeostasis in humans. The protein is believed to contain one or more of the following features: N-glycosylation site (PS00001); Glycosaminoglycan attachment site (PS00002); 4 cAMP- and cGMP-dependent protein kinase phosphorylation sites (PS00004); 12 Protein kinase C phosphorylation sites (PS00005); 8 Casein kinase II phosphorylation sites (PS00006); 8 N-myristoylation sites (PS00008); Zinc finger C2H2 type domain (PS50157); Cysteine-rich region (PS50311).

For SEQ ID NO: 3, residues 1-400 represent the 5'-UTR (untranslated region) while the 3'UTR comprises residues 1952-end. For the mutant sequence of SEQ ID NO: 6, showing the T1802G mutation, the 5'-UTR region is shown as residues 1-400 and the 3'-UTR region as residues 1952-end. For the mutant sequence of SEQ ID NO: 8, showing the 1706C1707 mutation, the 5'-UTR region is shown as residues 1-400 and the 3'-UTR region as residues 1952-end.

Gene95 has a Rhodanese-like domain most closely related to group of hypothetical, uncharacterized bacterial proteins. An examplary rhodanese-like domain is found at amino acids 304 - 406. There is no similarity of this gene to any other human gene, other than hypothetical or "rhodanese-like" or "sulfurtransferase-related" proteins. Rhodanese domains are found in a thiosulfate sulfurtransferase (e.g. Rhodanese; IPR001307). A public database search identified 21 hits in human genome with Rhodanese domains, 1 being a platelet adhesion molecule, the others being phosphatases. Other domain signatures such as a DnaJ central domain (e.g. [CXXCXGXG]4; amino acids 422-480), XS Zinc finger domain (e.g. amino acids 458-465) and integrase site (e.g. amino acids 452- 465) are also suggested.

Figure 5 sets out the nucleotide and amino acid sequences of individual exons, and some sequence from the promoter (genomic) and intron

regions (especially splice site junctions). The Gene95 sequence spans about 34kb of genomic DNA. Introns which are not fully set out may be obtained from the public human genome sequence database Goldenpath. Lower case letters indicate intron and promoter regions, while upper case letters indicate nucleotides in the cDNA or mRNA or amino acids, as dictated by the context. Exon 1 and Exon 2 have one or more asterisks ("*"). These are putative stop codons. The true initiator methionine is believed to be the first methionine in Exon 2 following the asterisk. Asterisks are also found in exon 10. The first asterisk is the expected termination site, generating a polypeptide of 516 amino acids. The remaining amino acid translations are illustrated for the sake of completeness and are not believed to lead to a functioning protein. Exon 6a and Exon 9a contain termination codons. If these alternate exons are used, they are believed to result in the truncated proteins, as indicated.

Homology: Among full length proteins, the closest homologs to human Gene95 are the shared 29 to 34 % identity of Rhodanese-related sulfurtransferases (Salmonella typhimurium and Brucella melitensis) and dihydrofolate reductase (Schizosaccharomyces pombe). Hypothetical bacterial proteins of highest homology include: Q9RVC9_DEIRA/119-216; Q55613_SYNY3/113-210; Q9Z7H1_CHLPN; 084632_CHLPN/113-212; 084632_CHLTR/113-212; 034131_LACLA/116-213; 031457_BACSU/116-213 (YBFQ protein).

This invention also discloses and claims mouse gene95 protein (SEQ ID NO: 5), the first non-human mammalian homolog of gene95. The sequence shares approximately 76% sequence identity across the full length of the protein with human gene95. Human and mouse amino acid sequences are aligned in Figure 6. Shared amino acids are indicated, along with conservative and non-conservative changes.

Additional investigations of short nucleic acid sequence segments from other mammals indicate that at the nucleotide level, mammalian species share a generally conserved gene95 sequence. Specifically:

Human:Cow 514/565 (90%) identical based on overlapping sequence
Human:Rat 338/401 (84%) identical based on overlapping sequence
Human:Pig 228/265 (86%),identical based on overlapping sequence
Human:Mouse 266/307 (86%) identical based on overlapping sequence

Expression pattern: Based on EST evidence, the gene is shown to be expressed in the following tissues and cell lines: teratocarcinoma, brain, placenta, ovarian tumor, lymphoma, melanoma, neuroblastoma. The data of Examples 2 and 3 confirm expression in a wide range of tissues.

Mutation 1, found in the family BC-11M, was identified as a 1802T>G (SEQ ID NO: 6) causing a Ser468Ala mutation SEQ ID NO: 7. The mutation was present in 13 affected family members and not present in 10 unaffected family members. The mutation was not found in 198 control chromosomes, nor in 164 chromosomes from individual with high HDL. Both control cohorts were of similar ethnicity to the BC-11M family. The mutation may result in loss of a putative Protein Kinase A phosphorylation site (PKA- basic-basic-neutral-Ser). This mutation alters a serine residue that is conserved in the mouse protein.

Mutation 2, from an individual in the NL-003 family with HDL level <5th percentile comprises a 1706C1707 insertion SEQ ID NO: 8. The insert is in a second base of codon for Q436 resulting in Q436P and 39 additional unique amino acids producing a premature stop codon, 475aa total (SEQ ID NO: 9 and counting from the methionine at residue 14). Mutation 2 is also observed in an individual from family NL-120 with HDL <5th percentile. Mutation 2 was not observed in 228 control chromosomes of similar ethnicity.

Alignments of the amino acid sequence of mutation 1 and mutation 2 with the wild–type amino acid sequence of human gene95 is set out in Figure 7.

Single Nucleotide Polymorphisms of the coding region (cSNPs) of *gene95* and other polymorphisms have been identified and are set out in Table 1. These cSNPs result in the amino acid changes identified, if any. Other cSNPs and polymorphisms in exonic or intronic regions of gene95 can also be identified by those skilled in the art.

Because of the processing that may take place in transforming the initial RNA transcript into the final mRNA, the nucleotide sequences disclosed herein may represent less than the full genomic sequence. Such genes and cDNA sequences are still considered corresponding sequences because they both encode similar RNA sequences. Thus, by way of non-limiting example only, a gene that encodes an RNA transcript, which is then processed into a shorter mRNA, is deemed to encode both such RNAs and therefore encodes an RNA complementary to (using the usual Watson-Crick complementarity rules), or that would otherwise be encoded by, a cDNA (for example, a sequence as disclosed herein). Thus, the sequences disclosed herein correspond to genes contained in the cells and are used to determine relative levels of expression because they represent the same sequences or are complementary to RNAs encoded by these genes. Such genes also include different alleles and splice variants that may occur in the cells used in the processes of the invention.

Thus, the polynucleotides, such as the gene sequences disclosed herein, for use in the screening assays of the invention "correspond to" *gene95* if the polynucleotide encodes an RNA (processed or unprocessed, including naturally occurring splice variants and alleles) that is at least 50% identical, preferably at least 70%, 80%, 90% identical, and more preferably at least 95%, or 98% identical to, and especially having the sequence of, an

RNA that would be encoded by, or be complementary to, such as by hybridization with, a polynucleotide having SEQ ID NO: 3. In addition, genes including sequences at least 50% identical to SEQ ID NO: 3, preferably at least about 70%, 80%, 90% identical to such a sequence, more preferably at least about 98% identical to such sequence and most preferably comprising such sequences as are specifically contemplated by all of the processes of the present invention as being genes that correspond to these sequences. In addition, sequences encoding the same polypeptides and proteins as any of these sequences, regardless of the percent identity of such sequences, are also specifically contemplated by any of the methods of the present invention that rely on any or all of said sequences, regardless of how they are otherwise described or limited. Thus, any such sequences are available for use in carrying out any of the methods disclosed according to the invention. Such sequences also include any open reading frames, as defined herein, present within *gene95*. Such sequences include the mutant *gene95* sequences of SEQ ID NO: 6 and SEQ ID NO: 8, and sequences containing the alternative exons shown in Figure 5.

In accordance with the forgoing, the present invention encompasses an isolated polynucleotide comprising a polynucleotide sequence or the full complement of this polynucleotide sequence, wherein the polynucleotide sequence is at least 65%, preferably at least 75%, more preferably at least 85%, especially at least 90%, most preferably 95%, or even 98%, identical to SEQ ID NO: 3, with the sequence of SEQ ID NO: 3 especially preferred. The invention also includes an isolated polynucleotide comprising a polynucleotide sequence that encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO: 4.

Relating to the gene95 protein (SEQ ID NO: 4), the invention includes all those proteins, polypeptides and amino acid sequences corresponding to SEQ ID NO: 4. Such sequences corresponding to SEQ ID NO: 4 include any full length protein or fragment thereof which retains a biological activity of

gene95 protein. Such sequences are preferably at least about 50% identical to SEQ ID NO: 4, more preferably about 60%, 70%, 80% and most preferably at least about 90%, or 95%, identical to SEQ ID NO: 4. Such sequences include any ortholog or homolog of gene95 protein from any organism, preferably a eukaryote, more preferably a mammal, most preferably a rodent or human; and any amino acid sequence having at least about 50%, 60%, 70%, 80% and most preferably at least about 90% sequence identity thereto. Such sequences include mutant sequences SEQ ID NO: 7 and SEQ ID NO: 9, sequences corresponding to alternative open reading frames of *gene95* nucleotide sequences, and sequences found in the alternatively spliced exons in Figure 5.

In accordance with the foregoing, the present invention includes an isolated polypeptide comprising an amino acid sequence having at least 65%, preferably at least 75%, more preferably at least 85%, especially at least 90%, most preferably 95%, or even 98%, identity with the amino acid sequence set forth in SEQ ID NO: 4, wherein a polypeptide comprising the sequence of SEQ ID NO: 4 is especially preferred.

The present invention also relates to nucleic acid vector comprising such an isolated polynucleotide and a recombinant host cell comprising such a vector. The present invention also encompasses a method for producing the polypeptide of SEQ ID NO: 4 comprising culturing the host cell of claim 49 under conditions supporting production of the polypeptide, methods of which are all well known to those skilled in the art in light of the disclosure set forth herein.

Screening Assays using Gene95

The present invention provides screening assays using gene95 gene and/or protein for use in identifying therapeutic agents, especially agents

antagonistic low HDL levels. One such protocol involves the screening of chemical agents for ability to modulate, either upward or downward, the activity of gene95 so as to elevate HDL activity, such as by increasing HDL levels, especially plasma levels, in an animal, including humans. Agents that alter the activity of the *gene95* gene, or active fragments or portions of it, or that may modulate the activity of polypeptides encoded by *gene95*, or polypeptides that act as transcription factors to modulate the activity of such genes, or other related gene segments, such as enhancers or other regulatory genetic elements that modulate the activity of *gene95*, acting either in cis or trans fashion, are thereby identified and may prove useful in preventing, treating, or otherwise ameliorating, low HDL levels, or the effects of low HDL levels or diseases related to low HDL levels, such as by ameliorating reverse cholesterol transport or other lipid related processes, and/or for diseases such as cardiovascular disease, atherosclerosis, and other diseases.

In accordance with the invention disclosed herein, there are provided compositions and methods for treating patients having low HDL-C, vascular disease, dyslipidemia, atherosclerosis or a related disorder, or for ameliorating reverse cholesterol transport in the body, by administering compounds that modulate biological activity or expression of gene95.

The present invention specifically contemplates screening assays, such as where a large number of compounds are to be screened for activity in modulating therapeutic target biological activity. As to all such assays as disclosed herein, such modulation may include either an increase or a decrease in therapeutic target biological activity.

Those skilled in the art are able to identify measurable biological activities of a therapeutic target which can be usefully incorporated into low or high throughput screening assays. Non-limiting examples of such assays are described herein for illustration purposes. Based on these teachings, other

embodiments [redacted] therapeutic target screening assays will be directly reduced to practice.

The invention also provides assays for the identification of such therapeutic compounds and their analogs. Compounds that modulate the biological activity of gene95 gene or protein are considered useful in the invention. Exemplary screening assays (*in vitro* or *in vivo)* and computational assays (*in silico*) for the identification of such compounds are detailed below. The screening assays of the invention simplify the evaluation, identification and development of therapeutic agents for the treatment and prevention of low HDL and/or vascular diseases, dyslipidemias or related disorders. In general, the screening methods provide a facile means for selecting natural product extracts or synthetic compounds of interest from a large population (i.e. a chemical library) which are further evaluated and condensed to a few active core structures. Multiple analogs of such core structures may be developed and tested to identify those preferred analogs which have improved characteristics as therapeutic agents.

The present invention also relates to a method for identifying an agent that modulates the activity of a polypeptide whose activity alters high density lipoprotein (HDL)-activity, comprising:

a) contacting a candidate compound with a polypeptide encoded by a polynucleotide corresponding to *gene95* under conditions facilitating such activity; and

b) determining a change in the biological activity of said polypeptide as a result of said contacting;

wherein said change in activity identifies the test compound as an agent that modulates said polypeptide biological activity.

In a one embodiment of this process, the change in biological activity in step (b) is a decrease or increase in activity of the polypeptide. In a preferred embodiment, the activity is measured by measuring the activity of an enzyme,

such as where the polypeptide itself has enzyme activity that can be directly or indirectly assayed. Such assays may be conducted *in vitro* or *in vivo*.

In a preferred embodiment, said polypeptide has the amino acid sequence of SEQ ID NO: 4; in an alternative embodiment, said polynucleotide has SEQ ID NO: 3.

In another embodiment, the polypeptide is present in a cell extract, a liposome or an intact cell. The present invention specifically contemplates embodiments in which the cell is a recombinant cell line containing a heterologous *gene95* construct. In other embodiments, said cell or extract is engineered by other than genetic engineering, such as where the activity of a polypeptide is to be enhanced and the cell has been engineered to contain, or have on its surface, said polypeptide but wherein the polypeptide is present due to physical insertion of the polypeptide into the membrane or cytoplasm of the cell and not through expression of a gene contained in the cell.

The present invention also relates to a method for identifying an agent that modulates a Gene95 activity, comprising:

a) contacting a test compound with a polypeptide encoded by a polynucleotide corresponding to *gene95* under conditions supporting an activity of said polypeptide; and

b) determining a change in the activity of the polypeptide as a result of said contacting;

wherein said change in activity identifies the test compound as an agent that modulates a Gene95 activity.

Such determined change in activity in step (b) may be a decrease or an increase in activity. In one embodiment, the activity is measured by measuring the activity of an enzyme. In other embodiments of such method, the polypeptide is present in a lipid bilayer, including where this lipid bilayer is part of a liposome. In addition, the polypeptide may be part of an intact cell, such

as where the intact cell is a cell that has been engineered to comprise said polypeptide, such as where the intact cell is a recombinant cell that has been genetically engineered to express said polypeptide, preferably where the cell does not express said polypeptide absent said engineering. In all cases, the intact cell is preferably a mammalian cell.

The present invention also relates to a method for identifying an HDL-enhancing agent, comprising administering to an animal an effective amount of an agent found to have modulating activity using any of the assays disclosed herein and detecting an increase in plasma HDL activity in said animal due to said administering thereby identifying an agent useful in enhancing HDL activity.

In a preferred embodiment, said animal exhibits low HDL activity prior to administering said agent. In another preferred embodiment, the HDL-enhancing activity is an increase in HDL level in said animal, most preferably an increase in plasma HDL level and preferably where said animal is a human patient.

The present invention additionally relates to a method for treating a low-HDL related disorder in an animal afflicted with said disorder comprising administering to said animal an effective amount of an agent found to have HDL-enhancing activity using any of the assays of the invention. In a preferred embodiment, the animal to be treated is a human patient. In an additional preferred embodiment, the disorder to be treated is selected from the group consisting of low HDL diseases, vascular diseases and dyslipidemias.

The agents identified according to the processes of the invention are also useful for treating or preventing coronary artery disease, regardless of the HDL status of the patient. For example, a patient with normal HDL levels who has a family history of coronary artery disease would still be advised to

take a therapeutic agent according to the invention in order to elevate HDL levels and further reduce the risk of coronary artery disease. Thus, the patient does not need to have a dyslipidemia in order to be eligible for treatment according to the invention.

In general, the screening methods of the invention involve screening any number of compounds for therapeutically active agents by employing any number of *in vitro* or *in vivo* experimental systems. Exemplary methods useful for the identification of such compounds are detailed herein.

The methods of the invention simplify the evaluation, identification and development of active agents for the treatment and prevention of low HDL levels as well as cardiovascular disease (CVD) and dyslipidemias. In general, the screening methods provide a ready means for selecting natural product extracts or compounds of interest from a large population which are further evaluated and condensed to a few active and selective materials. Positive candidates from this pool are then purified and evaluated in the methods of the invention to determine their HDL-raising, anti-CVD or anti-dyslipidemia activities or both. The positive candidates may be used directly as therapeutic agents, or they may provide informative structures for structure-activity relationship (SAR) analysis and development of further analogs which then become the preferred therapeutic agents.

<u>Assessment of Gene 95 as a sulfurtransferase responsible for HDL regulation.</u>

While the human genetics data establishes that the function of Gene 95 is to maintain HDL levels in humans, the specific mechanism of action of Gene 95 remains to be established. Without being bound to any theory for a mechanism of action, the inventors suggest that Gene 95 may function as a sulfurtransferase, or similar type enzyme. This is based on the comparison of

the Rhodanese domain found in the protein. Rhodanese domains may be classified as active for sulfurtransferase activity, active for phosphatase activity or catalytically inactive. Human and mouse G95 contain consensus for sulfurtransferase within the rhodanese domain. Possible consensus for phosphatase within rhodanese domain is found in the human but not conserved in mouse G95. In addition, one other human gene with this consensus has been annotated as a sulfurtransferase activity.

Several types of evidence connect these activities to HDL metabolism. For example, thioredoxin binding protein Txnip is known to be involved in lipid metabolism based on the hyperlipidemic mouse phenotype.

It is understood that screening assays employing a gene 95 sulfurtransferase biological activity may be designed using a known substrate for sulfurtransferases, such as cyanide, thioredoxin and dihydrolipoate. Alternatively, screening assays may be designed based on previously disclosed sulfurtransferase assays. In one colorimetric assay for thiocyanate production, the assay measures conversion of thiosulfate and cyanide to sulfite and thiocyanate. In an assay for NADPH oxidation by thioredoxin reductase, the assay measures conversion of thiosulfate and thioredoxin to sulfite and thioredoxinpersulfide. Compounds which agonize or antagonize these conversions via Gene 95 interaction are modulators of the invention.

Functional assays may be based on the activity of a polypeptide encoded by one or more of the polynucleotides disclosed herein as corresponding to *gene95*, including candidates with high homology thereto, such as polynucleotides at least 90% or 95% identical to such genes or polypeptides with high homology to the polypeptides encoded by said genes. Such assays may employ drug screening technology such as (but not limited to) the ability of various dyes to change color in response to changes in assay conditions resulting from the activity of the polypeptides.

Drug screening assays can also be based upon the ability of gene95 polypeptide to interact with other proteins. Such interacting proteins can be identified by a variety of methods known in the art, including, for example, radioimmunoprecipitation, co-immunoprecipitation, co-purification, and yeast two-hybrid screening. The ability of test compounds to agonize or antagonize protein-protein interactions is a standard type of screening assay that may be employed. Such interactions can be assayed by means including but not limited to fluorescence polarization or scintillation proximity methods. Drug screens can be based upon a function or feature apparent upon creation of a transgenic or knockout mouse, or upon overexpression of the protein or protein fragment in mammalian cells *in vitro*. Moreover, expression of mammalian (e.g., human) gene95 polypeptides in yeast or *C. elegans* allows for screening of candidate compounds in wild-type and mutant backgrounds, as well as screens for mutations that enhance or suppress a low HDL phenotype. Modifier screens can also be performed in transgenic or knock-out mice.

It will be evident to those skilled in the art that the disclosure herein facilitates incorporation of Gene95 into standard screening assays for identifying compounds that interact with and/or modulate gene95. Such assays include protein/compound binding assays, gene expression assays and many other types. Along with standard assays, known features of Gene95 suggest more specialized assays that may be employed. For example, the putative phosphorylation sites suggest that an assay which measures a test compound's propensity to influence the rate, amount or timing of phosphorylation of gene95 would be a useful modulator of the invention, and a potential therapeutic agent or analog thereof for the diseases indicated herein. Features of gene95 which can be exploited for use in screening include the N-glycosylation site, the Glycosaminoglycan attachment site; the 4 cAMP- and cGMP-dependent protein kinase phosphorylation sites; the 12 Protein kinase C phosphorylation sites; the 8 Casein kinase II

phosphorylation sites; the 8 N-myristoylation sites; the Zinc finger C2H2 type domain; and the Cysteine-rich region. Other features that may be useful for measurement, include the activity of the Rhodanese-like domain, the DnaJ central domain, the XS Zinc finger domain and the integrase site.

In accordance with the foregoing, the present invention relates to a method for identifying an agent that modulates the activity of a polynucleotide whose expression alters high density lipoprotein (HDL)-activity *in vivo*, comprising:

a) contacting a test compound with a polynucleotide corresponding to *gene95* under conditions facilitating expression of the polynucleotide; and

b) determining a change in the expression of the polynucleotide as a result of said contacting;

wherein said change in expression identifies the test compound as an agent that modulates the activity of a polynucleotide whose expression alters high density lipoprotein (HDL)-activity.

In specific embodiments of such a process, the change in expression in step (b) may be a decrease or increase in expression of said polynucleotide or gene.

In another aspect, the present invention relates to a method for identifying an agent that modulates Gene95 activity, comprising:

a) contacting a test compound with a genetic construct comprising a reporter gene operably linked to a Gene95 promoter under conditions supporting transcription of said reporter gene;

b) determining a change in transcription of the reporter gene as a result of said contacting

wherein a change in said transcription indicates that the test compound is an agent that modulates Gene95 activity.

Such determined changes may be an increase or decrease in the recited transcription and wherein transcription is determined by measuring the amount of an expression product encoded by said reporter gene, such expression product including either an RNA or a polypeptide. In preferred embodiments, the reporter gene may be present in a liposome or in an intact cells, such as a mammalian cell. Preferred embodiments of such methods include use of a mammalian Gene95 promoter, such as a human promoter, preferably that shown in SEQ ID NO: 15, or a mouse promoter, preferably the promoter sequence in SEQ ID NO: 14. In all cases, the reporter gene is commonly a gene whose expression is easily measured and will be other than Gene95 itself, although the latter is not excluded for use as a reporter gene.

Such a process is especially useful for identifying agents effective against low HDL levels and disorders, and symptoms thereof. Such agents or their chemical analogs are also useful for modulating gene95 gene or protein and/or for treatment or prophylaxis of cardiovascular disease. An example of a construct comprising a human Gene95 promoter sequence is shown in SEQ ID NO: 15, wherein the promoter sequence (about the first 1500 bases) is at the start of the sequence and is followed by exon 1 (starting at residue 1502) after which is intron 1 with exon 2 at the end of the sequence, while a sequence that incorporates a mouse promoter (about 1500 nucleotide residues) is shown in SEQ ID NO: 14, with the promoter at the beginning of the sequence, then followed by exon1 (starting at residue 1501), followed by intron 1 with exon 2 at the end of the sequence. The exonic sequences were found in mRNA. The present invention encompasses such exon and promoter sequences.

In another aspect, the invention provides a method for computationally identifying a compound for modulating HDL levels and/or treating a vascular disease or dyslipidemia. The method involves (a) determining the active site of the gene95 protein (i.e. through X-Ray crystallography or other techniques); and (b) through computational modeling, identifying a compound which

interacts with the active site, thereby identifying a compound, or its analog, as a compound which is useful for modulating HDL levels and/or treating a vascular disease, dyslipidemia or a related disorder.

The present invention further relates to a method for identifying a therapeutic agent, comprising:

a) contacting a chemical agent with a polynucleotide corresponding to *gene95* under conditions supporting expression of said polynucleotide;

b) determining a change in the expression of said polynucleotide as a result of said contacting;

wherein a change in said expression identifies a therapeutic agent.

In a further embodiment, the present invention encompasses a method for identifying a therapeutic agent that modulates the activity of a polypeptide that affects high density lipoprotein (HDL)-activity *in vivo*, comprising:

a) contacting a test compound with a polypeptide encoded by a polynucleotide corresponding to *gene95* under conditions supporting an activity of said polypeptide; and

b) determining a change in the activity of said polypeptide as a result of said contacting;

wherein a change in said activity identifies the test compound as a therapeutic agent that modulates the activity of a polypeptide that affects HDL activity.

In a preferred embodiment thereof, the polypeptide corresponds to SEQ ID NO: 4 and/or the polynucleotide corresponds to SEQ ID NO: 3

Gene95 polypeptide, preferably of mammals, most preferably of rodent or human, can be used as an antigen to raise antibodies, including monoclonal antibodies. Such antibodies will be useful for a wide variety of purposes, including but not limited to therapeutic uses for modulating gene95 biological activity, functional studies and the development of drug screening

assays and diagnostics. Monitoring the influence of agents (e.g., small organic compounds) on the expression or biological activity of gene95 polypeptide identified according to the invention can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase or decrease gene expression, protein levels, or biological activity can be monitored in clinical trails of subjects exhibiting low HDL levels due to inadequate gene expression, protein levels, or biological activity. Alternatively, the effectiveness of an agent determined by a screening assay to modulate expression of gene95, as well as structurally and functionally related genes, including genes with high homology thereto, and including protein levels, or biological activity can be monitored in clinical trials of subjects exhibiting decreased altered gene expression, protein levels, or biological activity. In such clinical trials, the expression or activity of the genes or polypeptides disclosed herein and, preferably, other genes that have been implicated in, for example, cardiovascular disease can be used to ascertain the effectiveness of a particular drug.

For example, and not by way of limitation, genes that are modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) that modulates the activity of gene95, or any expression products thereof, or polypeptides that modulate that activity of any of such genes (e.g., identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cholesterol levels or cardiovascular disease, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of gene95 and other genes implicated in similar or related disorders. The levels of gene expression can be quantified by Northern blot analysis or RT-PCR, or, alternatively, by measuring the amount of protein produced, by one of a number of methods known in the art, or by measuring the levels of biological activity of polypeptides encoded thereby or other genes. In this way, the gene expression can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this

response state may be determined before, and at various points during, treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, anti-gene95 antibody, antisense molecule, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of HDL activity or level of expression of gene95 protein, mRNA, or genomic DNA in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of gene95 protein, mRNA, or genomic DNA or HDL in the post-administration samples; (v) comparing the level of expression or activity of said protein, mRNA, or genomic DNA in the pre-administration sample with that of the corresponding post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of *gene95* or its encoded polypeptide, or to increase HDL-activity, such as plasma HDL-level, to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of the polypeptide or gene where excess expression contributes to lowering of HDL activity or levels.

The *gene95* genes disclosed herein as being involved in HDL regulation, such as inducing low HDL levels in an animal, can be used, or a fragment thereof can be used, as a tool to express a protein, where such genes encode a protein, in an appropriate cell *in vitro* or *in vivo* (gene therapy), or can be cloned into expression vectors which can be used to produce large enough amounts of protein to use in *in vitro* assays for drug screening. Expression systems which may be employed include baculovirus,

herpes virus, adenovirus, adeno-associated virus, bacterial systems, and eucaryotic systems such as CHO cells. Naked DNA and DNA-liposome complexes can also be used.

Assays of such activity include binding to intracellular interacting proteins, interaction with a protein that up-regulates gene or polypeptide activity, interaction with HDL particles or constituents, interaction with other proteins which facilitate interaction with HDL or its constituents, and measurement of cholesterol efflux. Furthermore, assays may be based upon the molecular dynamics of macromolecules, metabolites and ions by means of fluorescent-protein biosensors. Alternatively, the effect of candidate modulators on expression or activity may be measured at the level of protein production using the same general approach in combination with standard immunological detection techniques, such as Western blotting or immunoprecipitation with a specific antibody. Again, useful cholesterol-regulating or anti-CVD therapeutic modulators are identified as those which produce an change in activity that correlates with increased HDL levels, especially in the plasma of a test subject.

Candidate modulators (i.e., test compounds) may be purified (or substantially purified) molecules or may be one component of a mixture of compounds (*e.g.*, a combinatorial library or an extract or supernatant obtained from cells). In a mixed compound assay, gene expression is tested against progressively smaller subsets of the candidate compound pool (*e.g.*, produced by standard purification techniques, *e.g.*, HPLC or FPLC; Ausubel et al.) until a single compound or minimal compound mixture is demonstrated to modulate gene or protein activity or expression in a manner having a beneficial effect on HDL levels.

Specific compounds which will modulate the gene expression or gene transcript levels in a cell of gene95 include antisense nucleic acids, ribozymes and other nucleic acid compositions and sequence specific binding

compounds which specifically hybridize with said gene (including exons or introns of such genes) or any RNA transcript of *gene95*. These specific compounds are compounds of the invention, and are useful for treating the diseases discussed previously. Design and manufacturing of such compounds are easily achieved by those skilled in the art based on the disclosures of the instant specification.

Specific compounds which modulate the activity of gene95 polypeptide include antibodies (polyclonal or monoclonal) which specifically bind to an epitope of said polypeptide. These specific compounds are compounds of the invention, and are useful for treating the diseases discussed previously. Design and manufacturing of such compounds are well known to those skilled in the art.

Antibodies and antibody fragments that specifically recognize one or more epitopes of gene95, or epitopes of conserved variants of gene95, or peptide fragments of the gene95 are also encompassed by the invention. Such antibodies include but are not limited to polyclonal antibodies, monoclonal antibodies (MAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')2 fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. These may be used to treat any of the disorders disclosed herein.

As is known to those skilled in the art, use of fully humanized antibodies is generally preferred because they are generally less immunogenic and have a longer half-life when administered to humans. Leading available techniques include the UltiMab Human Antibody Development System(sm) featured by Medarex, Inc. (Princeton, NJ) which employs a transgenic mouse having a full suite of human genes for antibodies; and the Xenomouse technology of Abgenix, Inc. (Fremont, CA) which uses genetically engineered strains of mice in which mouse antibody

gene expression is suppressed and functionally replaced with human antibody gene expression, while leaving intact the rest of the mouse immune system.

These mice can be induced to generate fully humanized polyclonal antibodies, which are then converted to large-scale production using hybridoma technology. In one embodiment, the method includes immunizing a transgenic non-human animal, *e.g.,* a transgenic mouse, having a genome comprising a human heavy chain transgene and a human light chain transgene, with a purified or enriched preparation of the gene95 protein, or a fragment thereof. B cells (*e.g.,* splenic B cells) of the animal are then obtained and fused with myeloma cells to form immortal, hybridoma cells that secrete human monoclonal antiboides against the gene95 protein.

Humanized antibodies may also be produced, for example by replacing an immunogenic portion of an antibody with a corresponding, but non-immunogenic portion (i.e. chimeric antibodies) (Robinson, R. R. et al., International Patent Publication PCT/U.S.86/02269; Akira, K. et al., European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison, S. L. et al., European Patent Application 173,494; Neuberger, M. S. et al., PCT Application WO 86/01533; Cabilly, S. et al., European Patent Application 125,023; Better, M. et al., Science 240:1041-1043 (1988); Liu, A. Y. et al. Proc. Natl. Acad. Sci. USA 84:3439-3443 (1987); Liu, A. Y. et al., J. Immunol. 139:3521-3526 (1987); Sun, L. K. et al., Proc. Natl. Acad. Sci. USA 84:214-218 (1987); Nishimura, Y. et al., Canc. Res. 47:999-1005 (1987); Wood, C. R. et al., Nature 314:446-449 (1985)); Shaw et al., J. Natl. Cancer Inst. 80:1553-1559 (1988)). General reviews of "humanized" chimeric antibodies are provided by Morrison, S. L. (Science, 229:1202-1207 (1985)) and by Oi, V. T. et al., BioTechniques 4:214 (1986). Suitable "humanized" antibodies can alternatively be produced by CDR or CEA substitution (Jones, P. T. et al., Nature 321:552-525 (1986); Verhoeyan et al., Science 239:1534 (1988); Beidler, C. B. et al., J. Immunol. 141:4053-4060 (1988)).

For the production of antibodies, various host animals may be immunized by injection with the gene95, a gene95 peptide (e.g., one corresponding the a functional domain of the protein), truncated gene95 polypeptides (gene95 in which one or more domains has been deleted), functional equivalents of the gene95 or mutants of the gene95. Such host animals may include but are not limited to rabbits, mice, goats and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

There are various expression systems that can be used for the production of whole antibodies and antibody fragments. These include bacterial or mammalian cell culture and transgenic animals or plants. The expression system of choice is determined by the intended application and the desired yield, as is known by those skilled in the art. For example, animal cell culture and transgenic expression systems are desirable if glycosylation of the antibody is required, whereas bacterial expression systems are more efficient for production of unglycosylated antibodies, Fab fragments and the like. *See* Chad, HE and Chamow, SM. 2001. Curr. Opin. Biotech. 12: 188-193.

To generate fully human monoclonal antibodies to gene95, HuMab mice can be immunized with a preparation of gene95 or a fragment thereof, which has been purified or enriched according to any standard method. HuMab mice contain a human immunoglobulin gene miniloci that encodes

unrearrang[redacted]human heavy (mu and gamma) [redacted]d kappa light chain immunoglobulin sequences, together with targeted muattions that inactivate the endogenous mu and kappa chain loci (Lonberg, N. et al. (1994) Nature. 368(6474): 856-859.). These mice exhibit reduced expression of mouse IgM or kappa, and in response to immunization, the introduced human heavy and light chain transgenes undergo class switching and somatic mutation to generate high affinity human IgGkappa monoclonals.

Preferably mice are 6 – 16 weeks of age upon the first immunization. Initial immunization is intraperitoneal (IP) with antigen in complete Freund's adjuvant, followed by every other week IP immunizations (up to a total of 6) with antigen in incomplete Freund's adjuvant. The immune response can be monitored over the course of the immunization protocal with plasma samples obtained by retroorbital bleeds. The plasma can be screened, for example by ELISA or flow cytometry, and mice with sufficient titers of anti-gene95 human immunoglobulin can be used for fusions. Mice can be boosted intravenously with antigen 3 days before sacrifice and removal of the spleen. It is expected that 2-3 fusions for each antigen are needed for success.

For generation of hybrodiomas producing human monoclonal antibodies to gene95, the mouse splenocytes are first isolated and fused with PEG to a mouse myeloma cell line based upon standard protocols. The resulting hybridomas are then screened for the production of antigen-specific antibodies. For example, single cell suspensions of splenic lymphocytes from immunized mice are fused to one-sixth the number of P3X63-Ag8.653 nonsecreting mouse myeloma cells (ATCC, CRL 1580) with 50% PEG. Cells are plated at approximately 2 X 10^5 in flat bottom microtiter plate, followed by a two week incubation in selective medium containing 20% fetal calf serum, 18% "653" conditioned media, 5% origen (IGEN), 4 mM L-glutamine, 1 mM L-glutamine, 1 mM sodium pyruvate, 5 mM HEPES, 0.055 mM 2-mercaptoethanol, 50 units/ml penicillin, 50 mg/ml streptomycin, 50 mg/ml gentamycin and 1X HAT (Sigma; the HAT is added 24 hours after the fusion).

After two weeks, cells are cultured in medium in which the HAT is replaced with HT. Individual wells are then screened by ELISA for human anti-gene95 monoclonal IgM and IgG antibodies. Once extensive hybridoma growth occurs, medium is observed usually after 10-14 days. The antibody secreting hybridomas are replated, screened again, and if still positive for human IgG, anti-gene95 antibodies, can be subcloned at least twice by limiting dilution. The stable subclones are then cultured *in vitro* to generate small amounts of antibody in tissue culture medium for characterization.

To purify human anti-gene95 antibodies, selected hybridomas can be grown in two-litre spinner-flasks for monoclonal antibody purification. Supernatants can be filtered and concentrated before affinity chromatography with protein A-sepharose. (Pharmacia, Piscataway, NJ). Eluted IgG can be checked by gel electrophoresis and high performance liquid chromatography to ensure purity. The buffer solution can be exchanged into PBS, and the concentration can be determined by OD_{280} using 1.43 extinction co-efficient. The monolconal antibodies can be aliquoted and stored at –80°C until required for use according to the methods of the invention.

Another composition which can serve to increase HDL levels and/or treat the diseases discussed herein, is a nucleic acid construct comprising a *gene95* gene and combined into a gene-therapy delivery vehicle. Such vehicles may be simple aqueous buffers (for naked DNA delivery), viral based vehicles or non-viral (especially lipid-based) vehicles. These specific compounds are compounds of the invention, and are useful for treating the diseases discussed previously. Design and manufacturing of such compounds easily achieved by those skilled in the art based on the instant specification.

Agonists, antagonists, or mimetics found to be effective at modulating the level of cellular expression or activity may be confirmed as useful in animal models (for example, primates, mice, pigs, dogs, rabbits, or chickens).

For example, the compound may ameliorate the low HDL levels of mouse or chicken hypoalphalipoproteinemias.

A compound that promotes an increase in expression or activity of a gene or polypeptide that has the function of facilitating HDL production is considered particularly useful in the invention; such a molecule may be used, for example, as a therapeutic to increase the level or activity of HDL and thereby treat a low HDL condition in an animal (for example, a human).

The invention permits the identification of other HDL-related genes that may play a role in HDL-activity regulation and thus mutations in such genes may be important in studying other genes that are involved in HDL levels. Thus, for example, mutations that serve to stabilize gene95 proteins that otherwise contribute to elevating HDL levels have value. Such mutations can be incorporated into any protein therapy or gene therapy undertaken for the treatment of low HDL-C levels. Similarly, compounds that increase the stability of a wild-type gene95 polypeptide or decrease its catabolism may also be useful for the treatment of low HDL-C or any other condition resulting from low HDL levels. Mutant forms of gene95 identified herein are also useful for this purpose. Further mutations and compounds can be identified using the methods described herein.

In one such embodiment, cells expressing a gene95 polypeptide having a mutation are transiently metabolically labeled during translation and the half-life of the said polypeptide is determined using standard techniques. Mutations that increase the half-life of the polypeptide are ones that increase protein stability. These mutations can then be assessed for HDL-related biological activity. They can also be used to identify proteins that affect the stability of corresponding mRNA or other proteins. One can then assay for compounds that act on these factors or on the ability of these factors to bind such polypeptides.

In another example, cells expressing a wild-type gene95 polypeptide are transiently metabolically labeled during translation, contacted with a candidate compound, and the half-life of the polypeptide is determined using standard techniques. Compounds that increase the half-life of the polypeptide are useful compounds in the present invention.

In other embodiments, treatment with an agonist of the invention may be combined with any other HDL-raising or anti-CVD or anti-dyslipidemia therapies, or other conditions which require HDL management.

The assays of the invention comprise protein based assays. The polypeptide encoded by the *gene95* gene (purified or unpurified) can be used in an assay to determine its ability to bind another protein (including, but not limited to, proteins found to specifically interact with proteins known to be involved in HDL-activity. Some examples would include ABCA1 protein and apolipoprotein. The effect of a compound on that binding is then determined.

In another type of binding assay for which the polypeptides encoded by genes disclosed herein are useful, the gene95 polypeptide (or a polypeptide fragment thereof or an epitope-tagged form or fragment thereof) is harvested from a suitable source (e.g., from a prokaryotic expression system, eukaryotic cells, a cell-free system, or by immunoprecipitation from expressing cells). The polypeptide is then bound to a suitable support (e.g., nitrocellulose or an antibody or a metal agarose column in the case of, for example, a his-tagged form of said polypeptide). Binding to the support is preferably done under conditions that allow proteins associated with the polypeptide to remain associated with it. Such conditions may include use of buffers that minimize interference with protein-protein interactions. The binding step can be done in the presence and absence of compounds being tested for their ability to interfere with interactions between said polypeptide and other molecules. If desired, other proteins (e.g., a cell lysate) are added, and allowed time to associate with the polypeptide. The immobilized polypeptide is then washed

to remove proteins or other cell constituents that may be non-specifically associated with it the polypeptide or the support. The immobilized polypeptide is then dissociated from its support, and proteins bound to it are released (for example, by heating), or, alternatively, associated proteins are released from the polypeptide without releasing the latter polypeptide from the support. The released proteins and other cell constituents can be analyzed, for example, by SDS-PAGE gel electrophoresis, Western blotting and detection with specific antibodies, phosphoamino acid analysis, protease digestion, protein sequencing, or isoelectric focusing. Normal and mutant forms of such polypeptide can be employed in these assays to gain additional information about which part of the polypeptide a given factor is binding to. In addition, when incompletely purified polypeptide is employed, comparison of the normal and mutant forms of the protein can be used to help distinguish true binding proteins.

In a specific example of such an assay, such assay is performed using a purified or semipurified protein or other molecule that is known to interact with a polypeptide encoded by a polynucleotide corresponding to gene95. This assay may include the following steps.

1. Harvest the gene95 polypeptide and couple a suitable fluorescent label to it;

2. Label an interacting protein (or other molecule) with a second, different fluorescent label. Use dyes that will produce different quenching patterns when they are in close proximity to each other vs. when they are physically separate (i.e., dyes that quench each other when they are close together but fluoresce when they are not in close proximity);

3. Expose the interacting molecule to the immobilized polypeptide in the presence or absence of a compound being tested for its ability to interfere with an interaction between the two; and

4. Collect fluorescent readout data.

Another assay is includes Fluorescent Resonance Energy Transfer (FRET) assay. This assay can be performed as follows.

1. Provide the gene95 protein or a suitable polypeptide fragment thereof and couple a suitable FRET donor (e.g.,. nitro-benzoxadiazole (NBD)) to it;

2. Label an interacting protein (or other molecule) with a FRET acceptor (e.g., rhodamine);

3. Expose the acceptor-labeled interacting molecule to the donor-labeled polypeptide in the presence or absence of a compound being tested for its ability to interfere with an interaction between the two; and

4. Measure fluorescence resonance energy transfer.

Quenching and FRET assays are related. Either one can be applied in a given case, depending on which pair of fluorophores is used in the assay.

Gene95 may act by altering membrane permeability, such as the permeability of membranes to ions. Such activity may be assayed for using vesicles, such as liposomes or intact cells, wherein such structures comprise gene95 polypeptides of the invention, which polypeptides are expressed in such vesicle, preferably an intact cell, such as a mammalian recombinant cell, and the permeability of the membrane of the cell is determined in the presence or absence of such expression. In the same way, such permeability can then be assayed in the presence and absence of chemical agents known to modulate the activity of gene95. Thus, the utility of these agents in enhancing the activity of proteins known to affect such membrane transport can be readily determined. In the same way, the ability of these agents to affect the transport of other molecules, such as lipids, especially HDL, across such membranes is likewise readily determined.

In performing such assays, the test cell, such as the aforementioned mammalian recombinant cell expressing gene95, or a polynucleotide corresponding to such gene is loaded with a reporter molecule (such as a

fluorescent [redacted] indicator whose fluorescent properti[redacted]ange when it binds a particular ion) that can detect ions, or alternatively, the external medium is loaded with such a molecule. A molecule which exhibits differential properties when it is inside the vesicle compared to when it is outside the vesicle may be used. For example, a molecule that has quenching properties when it is at high concentration but not when it is at another low concentration would be suitable. Alternatively, the effect of transport of ions (Ca, K, Na, metals and the like) across the membrane can be measured if the effect is to quench or unquench a dye which has been loaded into the medium or into the cell. The movement of the charged molecule (either its ability to move or the kinetics of its movement) in the presence or absence of a compound being tested for its ability to affect this process can be determined.

In still another assay, uptake of radioactive isotopes into or out of a vesicle can be measured. The vesicles are separated from the extravesicular medium and the radioactivity in the vesicles and in the medium is quantitated and compared.

As already disclosed above, the present invention also relates to assays that may employ transcription factors for one or more of the genes disclosed herein. The association between the polynucleotide to be tested and the binding factor may be assessed by means of any system that discriminates between protein-bound and non-protein-bound DNA (e.g., a gel retardation assay). The effect of a compound on the binding of a factor to such DNA is assessed by means of such an assay. In addition to binding assays, *in vitro* assays in which the regulatory regions of the gene are linked to reporter genes can also be performed. In a preferred embodiment, such polynucleotide is one that corresponds to *gene95*, most preferably wherein said polynucleotide has SEQ ID NO. 3.

The invention includes recombinant cell lines containing a gene95 gene construct. Recombinant cell lines expressing the corresponding protein

are tested to identify a relevant biological activity of the protein that can be modulated by exposure to test compound(s). Compound(s) are systematically screened to evaluate whether they modulate the identified biological activity and those that effectively do so are then therapeutic agents, or analogs thereof, according to the invention.

Recombinant cell lines are also preferred for the preparation of purified protein, if a purified protein assay is desired. Those skilled in the art are capable of producing recombinant cell lines and extracting protein fractions containing highly purified proteins. These samples can be used in a variety of binding assays to identify compounds which interact with the proteins. Compounds that interact are therapeutic agents of the invention, or analogs thereof.

The invention also comprises the upstream untranslated regions and promoter regions of gene95. A segment of the promoter region is included in Figure 5. Larger portions of this promoter region can be now be identified by those skilled in the art, using this disclosure. The 5'UTR (untranslated region) is disclosed in SEQ ID NO. 3, 6, 8 and Figure 5. Such genomic or untranslated regions may be included in plasmids which are used in assays to identify compounds which modulate the expression of the identified gene. In one such assay, the upstream genomic region is ligated to a reporter gene, and incorporated into an expression plasmid. The plasmid is transfected into a cell, and the recombinant cell is exposed to test compound(s). Those compounds which increase or decrease the expression of the reporter gene are then modulators of the gene/protein, and are considered therapeutic agents of the invention.

The genes disclosed herein may also be involved in regulating cholesterol efflux. A transport-based assay for such activity can be performed *in vivo* or *in vitro*. For example, the assay may be based on any part of the reverse cholesterol transport process that is readily re-created in culture, such

as cholesterol or phospholipid efflux. Alternatively, the assay may be based on net cholesterol transport in a whole organism, as assessed by means of a labeled substance (such as cholesterol).

For high throughput screening, fluorescent lipids can be used to measure lipid efflux. For phospholipids, a fluorescent precursor, C6-NBD-phosphatidic acid, can be used. This lipid is taken up by cells and dephosphorylated by phosphatidic acid phosphohydrolase. The product, NBD-diglyceride, is then a precursor for synthesis of glycerophospholipids like phosphatidylcholine. The efflux of NBD-phosphatidylcholine can be monitored by detecting fluorescence resonance energy transfer (FRET) of the NBD to a suitable acceptor in the cell culture medium. Suitable acceptors include rhodamine-labeled phosphatidylethanolamine, a phospholipid that is not readily taken up by cells. The use of short-chain precursors obviates the requirement for the phospholipid transfer protein in the media. For cholesterol, NBD-cholesterol ester can be reconstituted into LDL. The LDL can efficiently deliver this lipid to cells via the LDL receptor pathway. The NBD-cholesterol esters are hydrolyzed in the lysosomes, resulting in NBD-cholesterol that can now be transported back to the plasma membrane and efflux from the cell. The efflux can be monitored by the aforementioned FRET assay in which NBD transfers its fluorescence resonance energy to the rhodamine-phosphatidylethanoline acceptor.

Animal models are also useful in the assays of the invention. Compounds identified as having activity in any of the above-described assays are subsequently screened in any available animal model system, including, but not limited to, pigs, rabbits, and chickens, to see if HDL levels are elevated as a result of such compounds being administered in an appropriately effective amount to said animal. Test compounds are administered to these animals according to standard methods. Test compounds may also be tested in mice bearing mutations in interfere with cholesterol transport. Additionally, compounds may be screened for their

ability to enhance an interaction between a polypeptide encoded by a polynucleotide corresponding to *gene95* and any HDL particle constituent such as ApoAI, ApoAII, or ApoE.

Such cholesterol efflux assay measures the ability of cells to transfer cholesterol to an extracellular acceptor molecule and are usually dependent on the presence of a transporter molecule, such as ABCA1. Thus, the genes disclosed herein may play a role in modulating the activity of ABCA1, such as where a polypeptide encoded by a polynucleotide disclosed herein binds to, and modulates, the activity of ABCA1 or other HDL-related protein. In this procedure, cells are loaded with radiolabeled cholesterol by any of several biochemical pathways (Marcil et al., Arterioscler. Thromb. Vasc. Biol. 19:159-169, 1999). Cholesterol efflux is then measured after incubation for various times (typically 0 to 24 hours) in the presence of HDL3 or purified ApoAI. Cholesterol efflux is determined as the percentage of total cholesterol in the culture medium after various times of incubation. Compounds that modulate cholesterol efflux in this assay may be acting on the gene95 protein target. Preferably, the cell is a recombinant cell line containing a gene95 construct.

The invention also comprises a transgenic animal, such as a mouse, that has had one or both alleles of gene95 inactivated (*e.g.*, by homologous recombination) or conversely, a mouse having one or more additional copies of gene95 or additional mutant copies of gene95. Such mice, including transgenic mice, represent a useful animal model for screening for compounds that raise HDL-C levels. Such mice are also useful for experimental purposes to determine the role of gene95 in living systems. Such an animal can be produced using standard techniques. In addition to the initial screening of test compounds, the animals having different versions of such genes are useful for further testing of efficacy and safety of drugs or agents first identified using one of the other screening methods described herein. Cells taken from the animal and placed in culture can also be exposed

to test compounds. HDL-C levels can be measured using standard techniques, such as those described herein.

Compounds first identified as useful in elevating HDL activity using one or more of the assays of the invention may be administered with a pharmaceutically-acceptable diluent, carrier, or excipient, in unit dosage form. Conventional pharmaceutical practice may be employed to provide suitable formulations or compositions to administer such compositions to patients. Although oral administration is preferred, any appropriate route of administration may be employed, for example, parenteral, intravenous, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, aerosol, or oral administration. Therapeutic formulations may be in the form of liquid solutions or suspension; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.

Typically, a successful gene95 modulating therapeutic agent will meet some or all of the following criteria. Oral availability should be at or above 20% (so that at least 20% is absorbed into the blood stream). Animal model efficacy should be less than about 2 mg/kg and the target human dose between 50 and 250 mg/70 kg, although doses outside of this range may be acceptable. The therapeutic index (or ratio of toxic dose to therapeutic dose) should be greater than 100. The potency (as expressed by IC_{50} value) should be less than about 10 μM, preferably below 1 μM, and most preferably below 100 nM. The agent should show reversible binding (i.e., competitive inhibition) and should be reasonably selective over unrelated targets, or related targets that cause unwanted side effects. The required dosage should be no more than about once or twice a day or at meal times.

Methods well known in the art for making formulations are found in, for example, Remington: The Science and Practice of Pharmacy, (19th ed.) ed.

A.R. Gennaro AR., 1995, Mack Publishing Company, Easton, PA. Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated napthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for agonists of the invention include ethylenevinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, or example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.

In general, novel drugs for the treatment of aberrant cholesterol levels and/or CVD or dyslipidemia are identified from large libraries of both natural product or synthetic (or semi-synthetic) extracts or chemical libraries including combinatorial chemical libraries, according to methods known in the art. Those skilled in the field or drug discovery and development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the exemplary methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (*e.g.*, semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are

commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceangraphics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, *e.g.*, by standard extraction and fractionation methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and development readily understand that methods for dereplication (*e.g.*, taxonomic dereplication, biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of replicates or repeats of materials already known for their HDL-raising, anti-CVD and/or anti-dyslipidemia activities should be employed whenever possible.

When a crude extract is found to have cholesterol-modulating or anti-CVD activities, or both, further fractionation of the positive lead extract is necessary to isolate chemical constituent responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having cholesterol-modulating or anti-CVD activities. The same *in vivo* and *in vitro* assays described herein for the detection of activities in mixtures of compounds can be used to purify the active component and to test derivatives thereof. Methods of fractionation and purification of such heterogeneous extracts are known in the art. If desired, compounds shown to be useful agents for the treatment of pathogenicity are chemically modified according to methods known in the art. Compounds identified as being of therapeutic value are subsequently analyzed using any standard animal model known in the art.

The invention further comprises the use of gene95 protein or protein fragments as therapeutic agents, and includes any modified forms of gene95

which retain the biological activity of gene95. These therapeutic agents can be administered to patients in need thereof, such as patients who would benefit by increasing gene95 biological activity either locally, regionally or throughout the body, as administered. Such patients include, but are not limited to, patients suffering from low HDL, dyslipidemia, cardiovascular disease and other disorders.

Functional gene95 protein, and soluble derivatives thereof, such as peptides corresponding to functional domains of *gene95* or modified versions of gene95 protein, can be produced by standard methods known in the art. Administration to patients can be intravenous, oral, intraperitoneal, or by any other method known in the art.

Diagnostics and Pharmacogenomics with gene95

Gene expression, both comparable and absolute, as well as biological activity, and mutational analysis can each serve as a diagnostic tool for low HDL; thus determination of the genetic subtyping of the gene sequence can be used to subtype low HDL individuals or families to determine whether the low HDL phenotype is related to a given protein function, as with the families used in accordance with the present invention. This diagnostic process can lead to the tailoring of drug treatments according to patient genotype, including prediction of side effects upon administration of HDL increasing drugs (referred to herein as pharmacogenomics). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual is examined to determine the ability of the individual to respond to a particular agent).

The present invention further relates to processes diagnosing the presence of a disorder, especially at its earliest stages, or the risk to a patient of developing such disorders. Thus, in another aspect, the present invention

relates to a process for diagnosing the presence of disorder in a patient suspected of being afflicted therewith comprising detecting a mutation in the *gene95* gene in the genome of said patient. In a preferred embodiment, the present invention relates to such diagnostic processes where the mutation is the mutation in the nucleotide sequence as depicted in SEQ ID NO: 6 and/or SEQ ID NO: 8 and/or Table 1.

In particular embodiments of such a process, the mutation is detected in a sample of DNA taken from said patient. Such a sample may be obtained in any manner commonly used to obtain such sample, such as by blood sample, biopsy or other methods. Further in accordance with the processes of the invention, such detecting may be accomplished *in vivo*, such as where an *in situ* procedure is used to detect the presence of a mutated *gene95* gene or nucleotide sequence in a patient. The disorders that can be so detected are not limited to low HDL, dyslipidemia, cardiovascular disease, and the other disorders thereby included, disclosed previously

In accordance with the foregoing, the present invention relates to a process for diagnosing the presence of an gene95-linked disease in a patient suspected of being afflicted therewith comprising detecting a mutation in the gene95 gene in the genome or mRNA of said patient.

In preferred embodiments, the invention relates to such a process wherein said mutation is detected in a sample of DNA taken from the patient, most preferably wherein the mutation is the mutation of SEQ ID NO. 6 or SEQ ID NO. 8 or Table 1.

Many different techniques are known for detecting the presence or absence of such a mutation. In preferred embodiments, said detecting is accomplished by determining the ability of a nucleic acid probe comprising at least 15 contiguous nucleotides that are complementary to a mutated portion of the sequence of SEQ ID NO: 1, most preferably wherein said probe comprises at least 30 contiguous nucleotides, especially where the probe comprises at least 50 contiguous nucleotides.

The processes disclosed herein are also useful in determining the risk of a patient in developing one or more of the above disorders, so that the present invention also relates to a process for determining a patient's risk of developing a disorder where said patient is suspected to be at risk thereof, comprising detecting a mutation in the *gene95* gene in the genome or mRNA of said patient. Such detection may, in keeping with the foregoing, be by either *in vitro* or *in vivo* means. In a preferred embodiment, said disorder is low HDL, dyslipidemia and/or cardiovascular disease. In a preferred embodiment, the mutation is a mutation in the gene corresponding to SEQ ID NO: 1.

It is equally an embodiment of the invention to diagnose a disorder using detection of the gene95 protein in a tissue or body sample, especially the detection of mutant gene95 protein or the level of mutant or normal gene95 protein. Such detection may be achieved by a variety of methods, but typically involves use of an anti-gene95 antibody in a standard ELISA assay. On the basis of this disclosure, those skilled in the art may now develop diagnostic tools for detecting mutant or normal gene95 in patient samples, for the diagnosis of various disorders.

Agents, or modulators, having a stimulatory or inhibitory effect on the *gene95* gene, or gene product, can be administered to individuals to treat disorders associated with low HDL-activity or level. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in efficacy of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages

and therapeutic regimens. Accordingly, the genetic complement of an individual with respect to the genes disclosed herein can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons (Eichelbaum, M., Clin. Exp. Pharmacol. Physiol., 23:983-985, 1996; Linder, M. W., Clin. Chem., 43:254-266, 1997). In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). Altered drug action may occur in a patient having a polymorphism (e.g., an single nucleotide polymorphism or SNP) in promoter, intronic, or exonic sequences of *gene95*. Thus by determining the presence and prevalence of polymorphisms allow for prediction of a patient's response to a particular therapeutic agent. In particular, polymorphisms in the promoter region may be critical in determining the risk of HDL deficiency and CVD. In addition to mutations, polymorphisms of *gene95* may likewise be relevant to the identification and use of agents identified using the assays disclosed herein.

In addition, different agents may have different abilities to affect the genes of a signature gene set. For example, if a potential therapeutic agent, say, agent A, causes a gene, or group of genes, related to low HDL levels to exhibit decreased expression, such as where a lower amount of mRNA is expressed from said gene(s), or less protein is produced from said mRNA, but a second potential agent, say, agent B, while modulating the activity of the same or related genes causes said expression to be reduced to half, such as where only half as much mRNA is transcribed or only half as much protein is translated from said mRNA as for agent A, then agent B is considered to have twice as much therapeutic potential as agent A.

The present invention also relates to a process that comprises a method for producing a product, including the generation of test data, comprising identifying an agent according to one of the disclosed processes for identifying such an agent (i.e., the therapeutic agents identified according to the assay procedures disclosed herein) wherein said product is the data collected with respect to said agent as a result of said identification process, or assay, and wherein said data is sufficient to convey the chemical character and/or structure and/or properties of said agent. For example, the present invention specifically contemplates a situation whereby a user of an assay of the invention may use the assay to screen for compounds having a desired polypeptide, for example, enzyme, modulating activity, or gene expression modulating activity, and, having identified the compound, then conveys that information (i.e., information as to structure, dosage, etc) to another user who then utilizes the information to reproduce the agent and administer it for therapeutic or research purposes according to the invention. For example, the user of the assay (user 1) may screen a number of test compounds without knowing the structure or identity of the compounds (such as where a number of code numbers are used the first user is simply given samples labeled with said code numbers) and, after performing the screening process, using one or more assay processes of the present invention, then imparts to a second user (user 2), verbally or in writing or some equivalent fashion, sufficient information to identify the compounds having a particular modulating activity (for example, the code number with the corresponding results). This transmission of information from user 1 to user 2 is specifically contemplated by the present invention.

In accordance with the foregoing, the present invention relates to a method for producing test data with respect to the modulating activity of a test compound comprising:

(a) contacting a test compound with a Gene95 polynucleotide or Gene95 polypeptide, or a polynucleotide construct comprising a reporter gene

operably linked to a Gene95 promoter, under conditions wherein said polynucleotide or reporter gene is being transcribed or said polypeptide is active,

(b) determining a change in the activity of said polypeptide or transcription of said polynucleotide or reporter gene as a result of said contacting, and

(c) producing test data with respect to the modulating activity of said test compound based on a change in the transcription of the determined polynucleotide or reporter gene or activity of the determined polypeptide wherein said change shows modulating activity.

Example 1

Mutations in Gene95 cause low HDL in humans

METHODS

Patient Ascertainment and Sample Collection

Kindred BC11M, of Dutch origin was identified as part of a collaboration between the Department of Medical Genetics at the University of British Columbia and St Paul's Hospital in Vancouver, British Columbia. This kindred was ascertained based on a hypoalphalipoproteinemia phenotype (<5th percentile) in the absence of other lipid and clinical abnormalities (e.g. obesity and diabetes). A history of coronary artery disease was evident in this kindred at the time of ascertainment. Additional Dutch kindreds NL-003, NL-120, were ascertained as part of the patient referral program to the lipid clinic at the Academic Medical Center in Amsterdam based on abnormal cholesterol levels. A control cohort of normolipidemic individuals (HDL >20th percentile and <80th percentile) was used to validate the causal nature of the mutation.

Markers and Genotyping

Information regarding genetic markers was obtained from the Genome Database and GenBank. All markers used for genome scan and fine mapping had been previously localized on the Marshfield map. An 'x' was appended to many of the marker names for internal use. Genomic DNA was genotyped using an Applied Biosystems Prism 3100 Genetic Analyzer running GeneMapper software (Applied Biosystems). Mendelian inheritance of alleles was verified using the PedCheck program.

Linkage Analyses

Two-point linkage analysis was performed using the MLINK program v4.1p from the FASTLINK software package and multipoint linkage analysis was performed using VITESSE v.1.0. An autosomal dominant inheritance model was used, assuming that 10% of all low HDL (<5th percentile) in the population is due to a mutation in the gene. The disease allele frequency was assumed to be 0.25%, with 98% penetrance and a phenocopy rate of 4.5% in non-carriers. Marker allele frequencies were estimated from a large collection of American and European samples. Haplotypes were manually reconstructed and cross-validated using SIMWALKv2.8.

RESULTS

Linkage analysis of family BC11M:

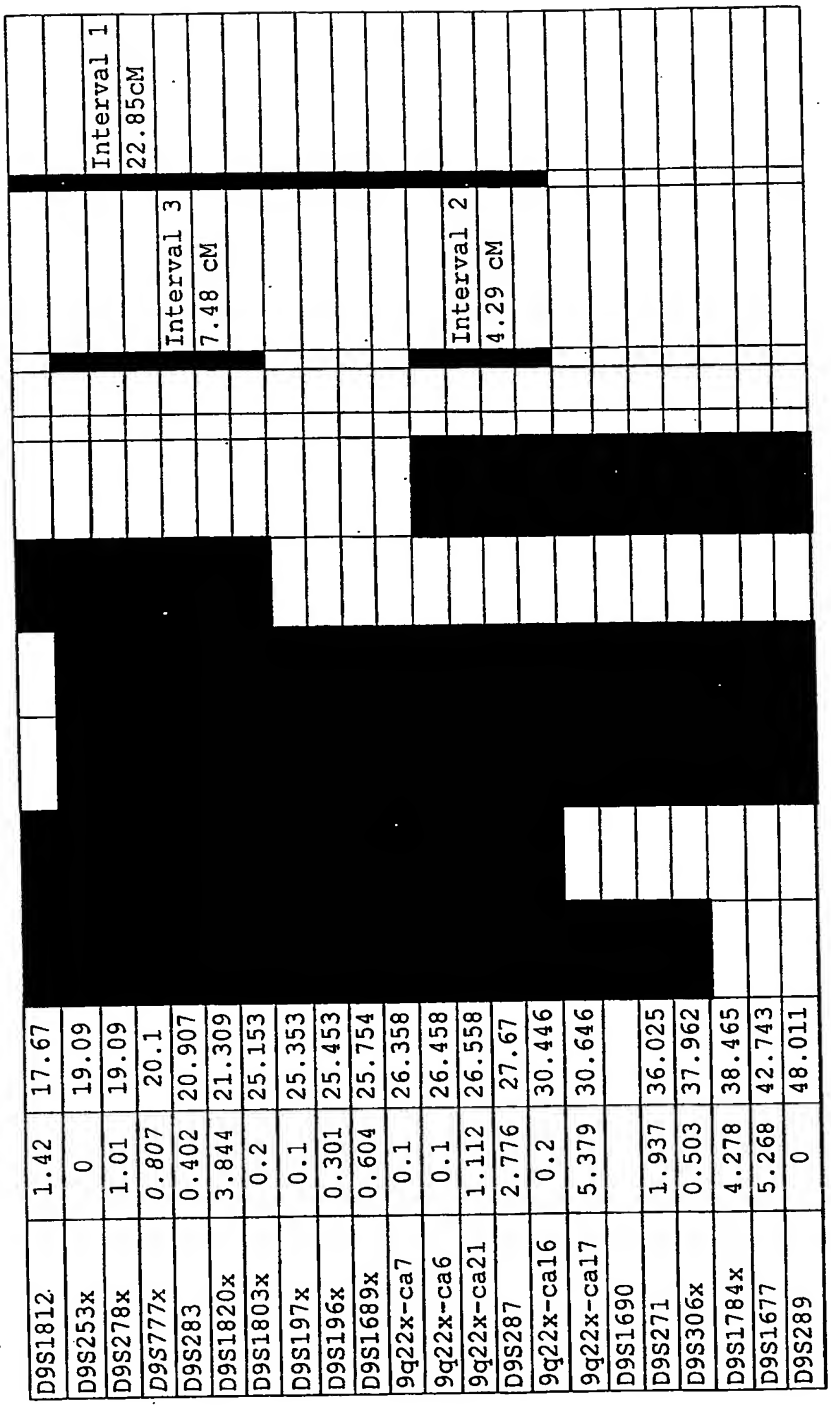
Twenty-four individuals from family BC11M were included in a 400 marker genome scan of 12 multiplex pedigrees segregating isolated hypoalphalipoproteinemia in a dominant fashion. Power analysis suggested that the BC11M kindred was suitable to identify an HDL locus using a linkage based approach under an autosomal dominant model (MaxE (Zmax 5cM) = 4.762). Two point linkage analysis of genome scan data gave a suggestive LOD score on chromosome 9 (D9S1677 LOD =2.285). This LOD score was followed up by multipoint linkage analysis, saturation genotyping and haplotyping with additional markers on an expanded pedigree. The region on

Table 2. Candidate gene position definitions based on affected and unaffected individuals in the BC11M pedigree. Two candidate gene intervals are possible when using unaffected individuals.

			BC11M Affecteds		BC11M Unaffected			
MARKER	cM	cM map	308	311	4141	4077	4055	4057
D9S273	0.662	0						
D9S166x	0.177	0.662						
D9S1799x	1.91	0.839						
D9S1876x	2.923	2.749						
D9S175	1.215	5.672						
D9S1834	0.908	6.887						
D9S1674	1.937	7.795						
D9S1780x	0.705	9.732						
D9S1843	3.2	10.437						
D9S264x	0.402	13.637						
D9S167	0.402	14.039						
D9S152x	0.301	14.441						
D9S1877x	1.523	14.742						
D9S1790x	0.2	16.265						
D9S1120x	0.1	16.465						
D9S1865x	0.301	16.565						
D9S249x	0.503	16.866						
D9S252x	0.301	17.369						

Table 2 – Continued

D9S1812	1.42	17.67
D9S253x	0	19.09
D9S278x	1.01	19.09
D9S777x	*0.807*	20.1
D9S283	0.402	20.907
D9S1820x	3.844	21.309
D9S1803x	0.2	25.153
D9S197x	0.1	25.353
D9S196x	0.301	25.453
D9S1689x	0.604	25.754
9q22x-ca7	0.1	26.358
9q22x-ca6	0.1	26.458
9q22x-ca21	1.112	26.558
D9S287	2.776	27.67
9q22x-ca16	0.2	30.446
9q22x-ca17	5.379	30.646
D9S1690		
D9S271	1.937	36.025
D9S306x	0.503	37.962
D9S1784x	4.278	38.465
D9S1677	5.268	42.743
D9S289	0	48.011



chromosome 9 was supported by multipoint linkage analysis, and the multipoint LOD rose to 8.0, defining an affecteds only interval of 20.91 cM (Interval 1; D9S1780 to 9q22xca17) (Table 2). This interval could be further into sub-regions based on chromosomal breakpoints in unaffected individuals (Interval 2; 9q22xca7 to 9q22x-ca17) (Interval 3; D9S1812 to D9S1803x) (Table 2). Interval 2 was the focus for the most comprehensive mutation detection effort.

Example 2

Tissue distribution of Gene 95 using an Origene cDNA panel.

Twenty-four serially diluted (4-log range to ensure linear amplification) cDNA clones arrayed into a multi-well PCR plate were obtained from Origene (Rockville, MD). PCR was performed using gene-specific primers for Gene 95 (sometimes "G95" herein). Reactions were in 10 µl and contained 1.0 µl 10x buffer (Roche Molecular Biochemicals, Indianapolis, IN, USA), 0.25 mM of each dNTP (Boehringer Mannheim), 5 pmole of forward primer and reverse primer, and 0.25 unit of Taq DNA polymerase (Roche Molecular Biochemicals, Indianapolis, IN, USA).

Forward primer:	CTGTGTCAGCACCTGCACCTC	SEQ ID NO. 10
Reverse primer:	ATCCCCATGGGCACGATTTC	SEQ ID NO. 11

The amplified fragment lies between exons 4 and 6, and has a product size of 230 bp.

DNA Engine Tetrad cyclers (MJ Research, Inc., Waltham, MA, USA) were used for PCR amplification. The conditions for PCR were a 3-min denaturation step at 94°C, followed by 35 cycles at 94°C for 30 sec, 50°C for 30 sec, and 72°C for 1 min. A final extension step of 7 min at 72°C was performed and samples were maintained at 10°C. PCR products were

analyzed by electrophoresis on 2% agarose gels in 1x TBE. A 100 bp ladder (Gibco-BRL, Rockville, MD, USA) was used as size standard.

Table 3

Human Tissue (Origene)	G95 Tissue Expression Levels
Tissue	by PCR
Muscle	H
Heart	H
Placenta	H
Plasma Blood Leucocytes	H
Brain	M
Lung	M
Kidney	L
Liver	L
Spleen	L
Small Intestine	L
Adrenal Gland	H
Fetal Brain	H
Fetal Liver	H
Ovary	H
Prostate	H
Stomach	H
Testis	H
Thyroid	H
Pancreas	M
Salivary	M
Uterus	M
Bone Marrow	L
Skin	L
Colon	no expression

Relative levels of G95 cDNAs are set forth in Table XX. H=high, M=medium, L=low.

Example 3

Tissue Expression by Northern Blot

Quantitation and sizing of the Gene 95 transcript were analyzed by use of a human 12-Lane multiple tissue Northern blot (BD Biosciences Clontech, Palo Alto, CA, USA). The human blot was hybridized with a 401 bp probe for Gene 95 from exons 6 to 9. A 4.4 kb band was identified, as expected. Actin was used as the control.

Forward primer	5' TTTCCAGAATTGCGTGTTGGTG 3'	SEQ ID NO. 12
Reverse primer	5' TGGATGCCACCCTTGAGCTG 3'	SEQ ID NO. 13

Results (Figure 9) show that Gene 95 is ubiquitously expressed with highest levels in muscle and heart. These results were confirmed by Northern blot analysis of mouse tissues (data not shown). In one such run, using fibroblasts, approximately 24 hours prior to cholesterol treatment, 2×10^5 cells were plated into each well in a 6-well format so as to obtain 50-70% confluency the following day. Cells were treated with a final concentration of 30 μg/ml cholesterol for 24 h followed by total RNA preparation. The control was Gene 95 expression from unaffected cell lines.

Example 4
Gene 95 expression in patient fibroblasts

The Ser to Ala mutation within exon 10 of Gene 95 is present in four patient fibroblasts (400, 558, 549, and 575). We decided to compare the expression of Gene 95 in these fibroblasts to normal controls (226, 475, and 485). This analysis was performed in the absence and presence of exogenous cholesterol, in order to explore a connection between lipid metabolism and Gene 95. Quantitative Real Time PCR (or "TaqMan") assays

were performed as follows. Human Gene 95 primers and their probes were designed using Primer Express software (Applied Biosystems, Foster City, CA).

Forward primer: 5' CCAAGGGAGTGTGCAAGG -3' SEQ ID NO. 31
Reverse primer: 5' CCTTTGTAAAAGCCATCAGGAAACT-3' SEQ ID NO. 32

RT-PCR was carried out in an ABI 7900 sequence detection system in a final volume of 25 ul, containing 120 ng of total RNA for human Gene 95 or 50 ng for human apoA1, 200μM primers, and 600 μM probe in 1x Taqman one-step RT-PCR Master mix (Applied Biosystem, CA). GAPDH or 18s (human) (Applied Biosystems) was used as the internal control. Data quantification and analysis was performed according to manufacturer's protocol. Briefly, ΔC_T for the calibrator and samples was obtained by subtracting the VIC C_T value of 18s RNA or GAPDH from the FAM C_T value of the gene of interest. $\Delta\Delta C_T$ was calculated by subtracting the average ΔC_T (calibrator) values from ΔC_T (sample). The mRNA quantity for the calibrator is expressed as 1 x sample and all other quantities are expressed as a number of percentage changes relative to the calibrator. Each sample was assayed in triplicate and standard error (SE) was calculated. RT-PCR reactions without RNA, and reactions with RNA but without reverse transciptase were used as negative controls. The size of the PCR product was confirmed by agarose gel.

Overall, extremely low levels of expression of Gene 95 were found in all fibroblasts, including controls. TaqMan results indicated that Gene 95 expression in two patient cell lines (400 and 575) was down regulated by approximately 50% compared to the controls either with or without cholesterol treatment. Expression of Gene 95 in patient 558 did not vary significantly from the controls. However, the expression of Gene 95 in 549 was elevated above controls (data not shown).

A trend towards reduced Gene 95 expression in mutation carriers was noted. Messenger RNA expression levels may be sensitive to cell culture conditions, cell density and passage number. Without desiring to be bound to any theory of mechanism of action, it is suggested that this trend towards decreased levels of Gene 95 mRNA in mutation carriers may be the underlying mechanism of the low HDL condition in these patients.

WHAT IS CLAIMED IS:

1. A method for identifying an agent that modulates Gene95 activity, comprising:

a) contacting a test compound with a genetic construct comprising a reporter gene operably linked to a Gene95 promoter under conditions supporting transcription of said reporter gene;

b) determining a change in transcription of the reporter gene as a result of said contacting

wherein a change in said transcription indicates that the test compound is an agent that modulates Gene95 activity.

2. The method of claim 1 wherein the determined change in transcription of step (b) is a decrease in transcription.

3. The method of claim 1 wherein the determined change in transcription of step (b) is an increase in transcription.

4. The method of claim 1 wherein transcription is determined by measuring the amount of an expression product encoded by said reporter gene.

5. The method of claim 4 wherein the expression product is an RNA.

6. The method of claim 4 wherein the expression product is a polypeptide.

7. The method of claim 1 wherein the reporter gene is in a liposome.

8. The method of claim 1 wherein the reporter gene is in an intact cell.

9. The method of claim 8 wherein the intact cell is a mammalian cell.

10. The method of claim 1 wherein the promoter is a mammalian Gene95 promoter.

11. The method of claim 10 wherein the Gene95 promoter is a human promoter.

12. The method of claim 11 wherein the human promoter comprises the promoter sequence in SEQ ID NO: 15.

13. The method of claim 10 wherein the Gene95 promoter is a mouse promoter.

14. The method of claim 13 wherein mouse promoter comprises the promoter sequence in SEQ ID NO: 14.

15. The method of claim 1 wherein the reporter gene is not Gene95.

16. A method for identifying an agent that modulates a Gene95 activity, comprising:

a) contacting a test compound with a polypeptide encoded by a polynucleotide corresponding to *gene95* under conditions supporting an activity of said polypeptide; and

b) determining a change in the activity of the polypeptide as a result of said contacting;

wherein said change in activity identifies the test compound as an agent that modulates a Gene95 activity.

17. The method of claim 16 wherein the determined change in activity in step (b) is a decrease in activity.

18. The method of claim 16 wherein the determined change in activity in step (b) is an increase in activity.

19. The method of claim 16 wherein the activity is measured by measuring the activity of an enzyme.

20. The method of claim 16 wherein the polypeptide is present in a lipid bilayer.

21. The method of claim 20 wherein the lipid bilayer is part of a liposome.

22. The method of claim 16 wherein the polypeptide is part of an intact cell.

23. The method of claim 22 wherein the intact cell is a cell that has been engineered to comprise said polypeptide.

24. The method of claim 22 wherein the intact cell is a recombinant cell that has been genetically engineered to express said polypeptide.

25. The process of claim 24 wherein the cell does not express said polypeptide absent said engineering.

26. The process of claim 22 wherein said cell is a mammalian cell.

27. A method for identifying an HDL-enhancing agent, comprising administering to an animal an effective amount of an agent found to have modulating activity using an assay of claim 1, 13, 24 or 49 and detecting an increase in plasma HDL activity in said animal due to said administering thereby identifying an agent useful in enhancing HDL activity.

28. The method of claim 27 wherein the animal exhibits low HDL activity prior to administering said agent.

29. The method of claim 27 wherein the HDL-enhancing activity is an increase in HDL level in said animal.

30. The method of claim 29 wherein the increase in HDL level is an increase in plasma HDL level.

31. The method of claim 29 wherein the animal is a human patient.

32. A method for treating a low-HDL related disorder in an animal afflicted with said disorder comprising administering to said animal an effective amount of an agent found to have HDL-enhancing activity using the assay of claim 27.

33. The method of claim 32 wherein the animal is a human patient.

34. The method of claim 32 wherein the disorder is selected from the group consisting of low HDL diseases, vascular diseases and dyslipidemias.

35. A method for producing test data with respect to the modulating activity of a test compound comprising:

(a) contacting a test compound with a Gene95 polynucleotide or Gene95 polypeptide, or a polynucleotide construct comprising a reporter gene operably linked to a Gene95 promoter, under conditions wherein said polynucleotide or reporter gene is being transcribed or said polypeptide is active,

(b) determining a change in the activity of said polypeptide or transcription of said polynucleotide or reporter gene as a result of said contacting, and

(c) producing test data with respect to the modulating activity of said test compound based on a change in the transcription of the determined polynucleotide or reporter gene or activity of the determined polypeptide wherein said change shows modulating activity.

36. A method for treatment or prophylaxis of cardiovascular disease comprising administering to a patient in need thereof, a therapeutic or prophylactic dose of a compound which modulates gene95 protein biological activity or expression.

37. A method for identifying a therapeutic agent, comprising:

a) contacting a chemical agent with a polynucleotide corresponding to *gene95* under conditions supporting expression of said polynucleotide;

b) determining a change in the expression of said polynucleotide as a result of said contacting;

wherein a change in said expression identifies a therapeutic agent.

38. A method for identifying a therapeutic agent that modulates the activity of a polypeptide that affects high density lipoprotein (HDL)-activity *in vivo*, comprising:

a) contacting a test compound with a polypeptide encoded by a polynucleotide corresponding to *gene95* under conditions supporting an activity of said polypeptide; and

b) determining a change in the activity of said polypeptide as a result of said contacting;

wherein a change in said activity identifies the test compound as a therapeutic agent that modulates the activity of a polypeptide that affects HDL activity.

39. The method of claim 38 wherein said polypeptide corresponds to SEQ ID NO: 4.

40. The method of claim 38 wherein said polynucleotide corresponds to SEQ ID NO: 3

41. An isolated polynucleotide comprising a polynucleotide sequence or the full complement of the polynucleotide sequence, wherein the polynucleotide sequence is at least 95% identical to SEQ ID NO: 3.

42. An isolated polynucleotide comprising a polynucleotide sequence that encodes a polypeptide having the amino acid sequence set forth in SEQ ID NO: 4.

43. An isolated polynucleotide comprising a polynucleotide that has the sequence set forth in SEQ ID NO: 3.

44. An isolated polypeptide comprising an amino acid sequence having at least 95% identity with the amino acid sequence set forth in SEQ ID NO: 4.

45. The isolated polypeptide of claim 44, wherein the isolated polypeptide comprises an the amino acid sequence having at least 95% identity with the amino acid sequence set forth in SEQ ID NO: 4.

46. An isolated polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 4.

47. An isolated polypeptide consisting of the amino acid sequence set forth in SEQ ID NO: 4.

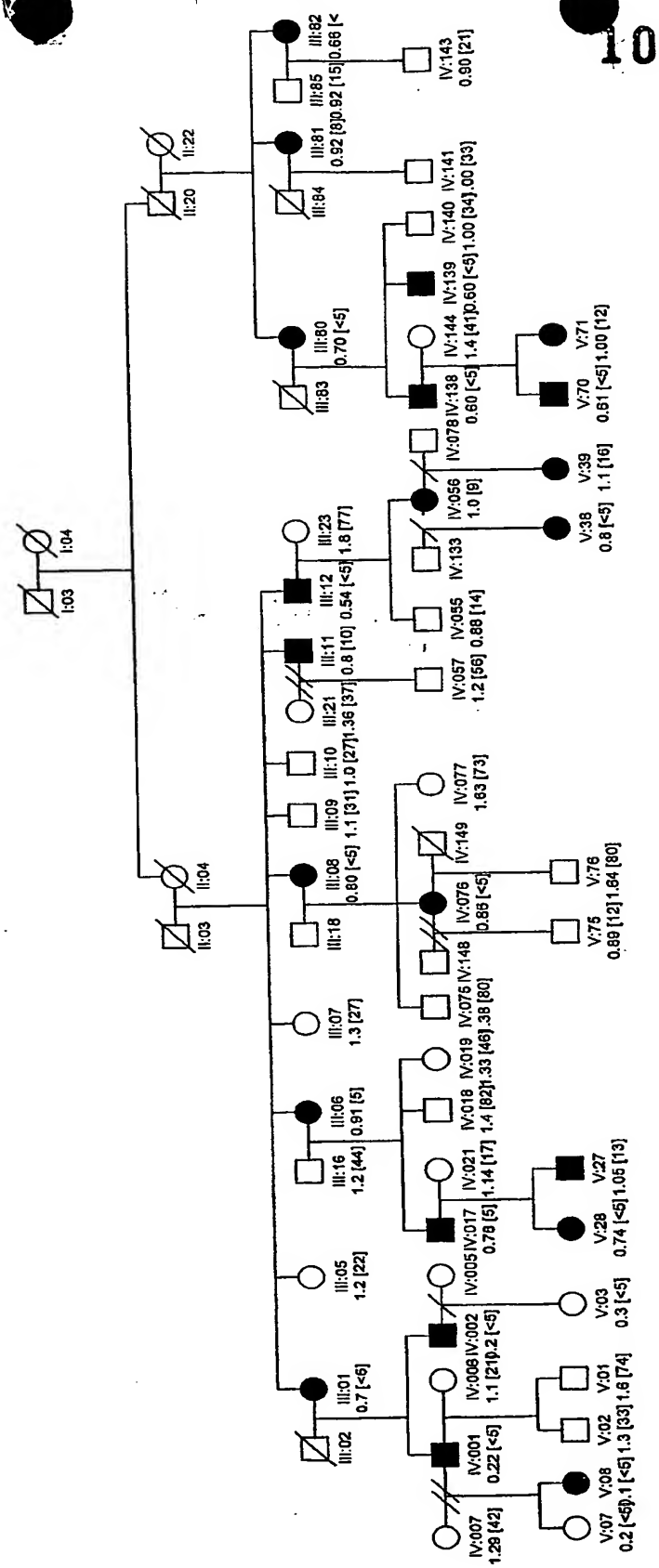
48. A nucleic acid vector comprising the isolated polynucleotide of claim 41.

49. A recombinant host cell comprising the vector of claim 48.

50. A method for producing the polypeptide of SEQ ID NO: 4 comprising culturing the host cell of claim 49 under conditions supporting production of the polypeptide.

Figure 1

BC-11M



NL-003

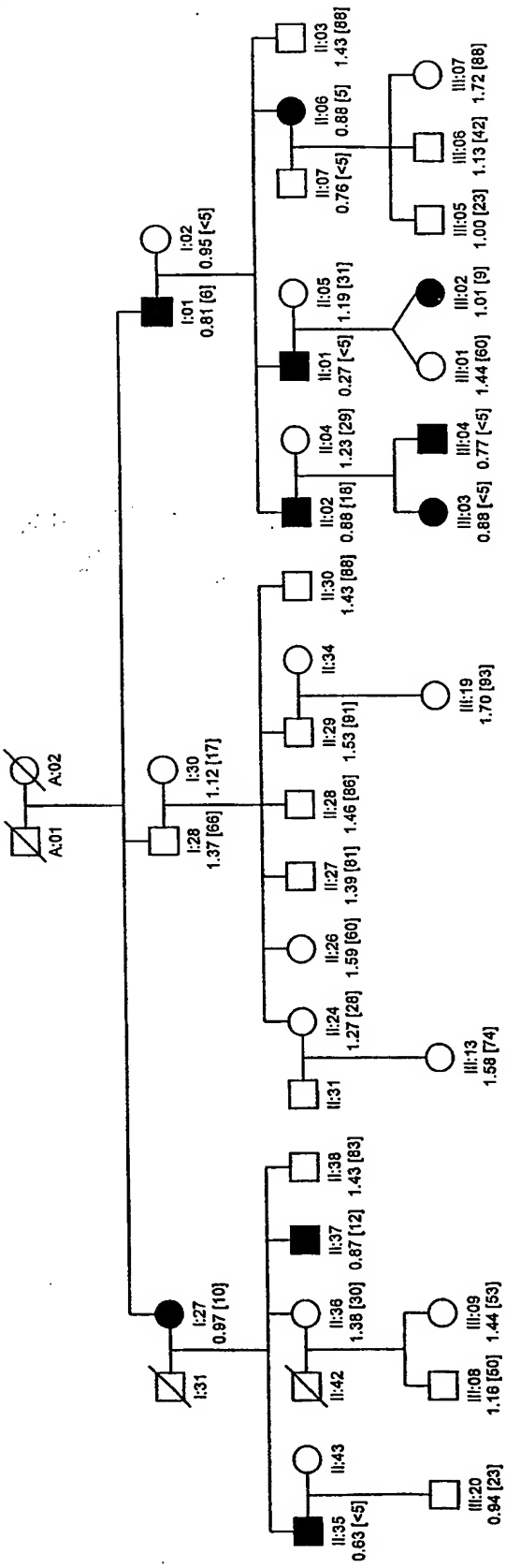


Figure 2

THIS PAGE BLANK (USPTO)

10/518868

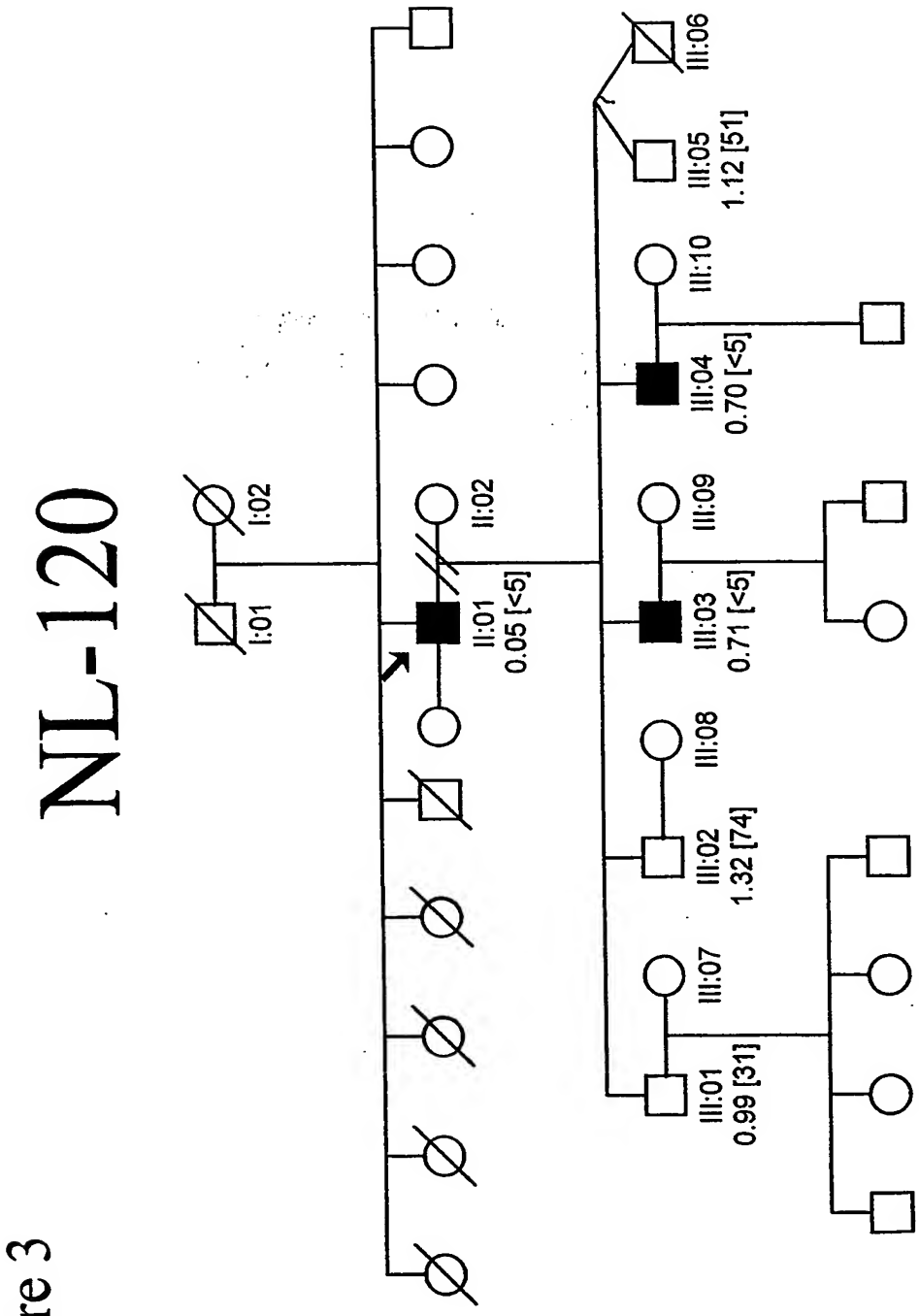


Figure 3

Fig. 4

G95 – Possible Transcripts Based on Bioinformatics

Figure 5A

```
aaggcgcgcggatgggccggcacgcaggcgcactagctcgccacggccccggaagcgcag
gagaggccgccgggtggggctaggcgctgcgacagccggcgtgaggaagctcagtgggct
acgaacgtctggcacacatgcaaccgccccctcgggctgcctccgcctgccggctacttc
tttctcccgccttccgctc

 E1: ATGTCCTTCCGGGGCCACGCAGAAAGTGCCGCCGCTTTGGCCACTCAGAGC
 +2:  C  P  S  G  A  T  Q  K  V  P  P  L  W  P  L  R  A
 E1: CCCCGGGCCGCGGTCGTCGTACGCCTGAAGGCGGGTCGTGCCGGCGGCCGC
 +2:  P  G  P  R  S  S  Y  A  *  R  R  V  V  P  A  A  A
 E1: TCTAGTCTCCGCCTCCGCTCAGGCCGGTCCTCCGGGGCTTCTCAATGGTTT
 +2:  L  V  S  A  S  A  Q  A  G  P  P  G  L  L  N  G  F
 E1: CCCGGTGGCCTCTCAATGGTTTTCCCGGCGGCCCTTGCGCCGACGCCAGGA
 +2:  P  V  A  S  Q  W  F  S  R  R  P  L  R  R  R  Q  E
 E1: GACTTCCGGAGCTTGGTGACGTCACGAGCGAGCTTTTCTACCCAAATACGC
 +2:  T  S  G  A  W  *  R  H  E  R  A  F  L  P  K  Y  A
 E1: GGCGGGGGAATAGGCTCGAGGGCGGTGAGCAGTGACAATTGCTAGGCGGAG
 +2:  A  G  E  *  A  R  G  R  *  A  V  T  I  A  R  R  R
 E1: ACAGTGCAGGGAAGAGAGACCTTAGAAAGGATCAGGACTGGCGG
 +2:  Q  C  R  E  E  R  P  *  K  G  S  G  L  A  G

                    gtatgt...cttatttatattctag

 E2: GAGGTATTTAACTGAAAGGAATATCTGCTTCACTGTTGCAACCAAACCAGA
 +3:   G  I  *  L  K  G  I  S  A  S  L  L  Q  P  N  Q  M
 E2: TGCCTTCTTCCACTTCACCAGACCAAGGAGATGACCTGGAGAACTGCATTT
 +3:   P  S  S  T  S  P  D  Q  G  D  D  L  E  N  C  I  L
 E2: TAAGATTTTCTGACCTGGATTTAAAAGATATGAGTCTTATTAATCCCAGCA
 +3:   R  F  S  D  L  D  L  K  D  M  S  L  I  N  P  S  S
 E2: GCAGTCTTAAAGCAGAATTAGATGGCAGTACAAAAAAGAAATACTCGTTTG0
 +3:   S  L  K  A  E  L  D  G  S  T  K  K  K  Y  S  F  A
 E2: CAAAGAAAAAG
 +3:   K  K  K

                    gtagaa...
```

THIS PAGE BLANK (USPTO)

Figure 5B


```
                ...tgaaatctaattgcag

E3: GCCTTTGCCCTTTTTGTCAAAACCAAAGAAGTTCCAACAAAAAGGAGTTTT
+1: A  F  A  L  F  V  K  T  K  E  V  P  T  K  R  S  F
E3: GAATGTAAAGAAAAATTGTGGAAATGCTGTCGGCAGCTATTCACAGACCAA
+1: E  C  K  E  K  L  W  K  C  C  R  Q  L  F  T  D  Q
E3: ACCAGCATCCATAGACATGTGGCAACACAACATGCTGATGAAATTTATCAC
+1: T  S  I  H  R  H  V  A  T  Q  H  A  D  E  I  Y  H
E3: CAGACAGCTTCTATTTTAAAGCAACTGGCTGTGACATTGAGCACCTCAAAG
+1: Q  T  A  S  I  L  K  Q  L  A  V  T  L  S  T  S  K
E3: AGTCTTTCGTCTGCAGATGAAAAGAACCCTTTAAAAGAGTGCCTTCCACAT
+1: S  L  S  S  A  D  E  K  N  P  L  K  E  C  L  P  H
E3: AGCCATGACGTGTCTGCTTGGCTCCCTGATATAAGCTGCTTTAACCCTGAT
+1: S  H  D  V  S  A  W  L  P  D  I  S  C  F  N  P  D
E3: GAGCTGATAAG
+1: E  L  I     S

          gtaaga...gattttcattttatag

E4: TGGCCAGGGCAGTGAAGAAGGGGAGGTGCTCCTTTATTACTGCTACCATGA
+2:  G  Q  G  S  E  E  G  E  V  L  L  Y  Y  C  Y  H  D
E4: CCTGGAGGATCCCCAATGGATCTGTGCCTGGCAGACAGCTCTGTGTCAGCA
+2:  L  E  D  P  Q  W  I  C  A  W  Q  T  A  L  C  Q  H
E4: CCTGCACCTCACAGGCAAG
+2:  L  H  L  T  G  K

                gtaaca...ccgtcttgtgtctcag

E5: ATTCGAATTGCTGCAGAAGGAATCAATGGGACAGTTGGTGGAAGCAAATTG
+1: I  R  I  A  A  E  G  I  N  G  T  V  G  G  S  K  L
E5: GCTACCAGACTTTATGTGGAAGTCATGCTTTCCTTCCCATTGTTTAAGGAT
+1: A  T  R  L  Y  V  E  V  M  L  S  F  P  L  F  K  D
E5: GACCTGTGTAAAGATGATTTTAAG
+1: D  L  C  K  D  D  F  K

                gtaaga...
```

Figure 5C

```
                ...gtttctcattggctag

E6: ACCAGCAAAGGAGGAGCTCACTGTTTTCCAGAATTGCGTGTTGGTGTATTT
+1: T  S  K  G  G  A  H  C  F  P  E  L  R  V  G  V  F
E6: GAAGAAATCGTGCCCATGGGGATCAGCCCCAAAAAGATCTCCTACAAGAAG
+1: E  E  I  V  P  M  G  I  S  P  K  K  I  S  Y  K  K
E6: CCTG
+1: P  G

          gtatgc...tttggtttggttttag

E7: GAATCCATTTATCCCCAGGTGAATTTCATAAAGAAGTAGAAAAGTTTTTAT
+3:   I  H  L  S  P  G  E  F  H  K  E  V  E  K  F  L  S
E7: CTCAGGCAAATCAAGAACAAAGTGATACTATCCTTCTTGATTGCAGAAACT
+3:   Q  A  N  Q  E  Q  S  D  T  I  L  L  D  C  R  N  F
E7: TCTATGAAAGCAAAATA
+3:   Y  E  S  K  I

          gtaagt...tgctcctatgttacag

E8: GGACGATTCCAAGGCTGCTTAGCCCCAGACATCAGGAAATTCAGTTACTTC
+1: G  R  F  Q  G  C  L  A  P  D  I  R  K  F  S  Y  F
E8: CCTAGCTACGTTGACAAAAATCTAGAACTTTTCAGAGAGAAGAGAGTGCTG
+1: P  S  Y  V  D  K  N  L  E  L  F  R  E  K  R  V  L
E8: ATGTACTGTACCGGGGGCATCCGCTGTGAGCGGGGTTCAGCCTACCTCAAA
+1: M  Y  C  T  G  G  I  R  C  E  R  G  S  A  Y  L  K
E8: GCCAAG
+1: A  K

          gtgagc...gtttttccacacctag

E9: GGAGTGTGCAAGGAGGTGTTCCAGCTCAAGGGTGGCATCCACAAGTACCTG
+1: G  V  C  K  E  V  F  Q  L  K  G  G  I  H  K  Y  L
E9: GAAGAGTTTCCTGATGGCTTTTACAAAGGGAAGTTGTTTGTTTTTGATGAA
+1: E  E  F  P  D  G  F  Y  K  G  K  L  F  V  F  D  E
E9: CGCTATGCTCTGTCCTACAACAGTGATGTGGTGTCAG
+1: R  Y  A  L  S  Y  N  S  D  V  V  S   E

          gtaggt...
```

Figure 5D

```
                    ...tttccttcctccccag

E10: AGTGTTCATACTGTGGAGCCCGCTGGGACCAGTATAAACTCTGCTCTACTC
 +3:   C  S  Y  C  G  A  R  W  D  Q  Y  K  L  C  S  T  P
E10: CCCAGTGCCGCCAGCTCGTTTTGACCTGCCCTGCCTGTCAAGGACAAGGAT
 +3:   Q  C  R  Q  L  V  L  T  C  P  A  C  Q  G  Q  G  F
E10: TCACAGCCTGTTGTGTCACATGTCAAGACAAGGGGAGCAGGAAAGTTTCAG
 +3:   T  A  C  C  V  T  C  Q  D  K  G  S  R  K  V  S  G
E10: GCCCTATGCAAGACAGCTTTAAAGAGGAATGCGAGTGCACAGCCCGACGGC
 +3:   P  M  Q  D  S  F  K  E  E  C  E  C  T  A  R  R  P
E10: CACGCATACCTAGGGAACTCTTGCAGCATGTGCGACAGCCTGTGAGCCCAG
 +3:   R  I  P  R  E  L  L  Q  H  V  R  Q  P  V  S  P  E
E10: AGCCAGGGCCTGATGCTGATGAGGATGGGCCAGTGCTTATGTGAGCAGCAC
 +3:   P  G  P  D  A  D  E  D  G  P  V  L  M  *  A  A  P
E10: CTTTGGCATTTTCCCAGGCCCTCGGTAAAAGTAGGTTTGGGGTGACTATAC
 +3:   L  A  F  S  Q  A  L  G  K  S  R  F  G  V  T  I  Q
E10: AGAGAAAGCATGGCAAGACTGCAGAAACAGAGAAATCGGGAACTTCAGTTC
 +3:   R  K  H  G  K  T  A  E  T  E  K  S  G  T  S  V  L
E10: TGGCCGCTGCCACCGTGGCAGCCGTCTACACTTCACAGCGGGAGGGGAGGA
 +3:   A  A  A  T  V  A  A  V  Y  T  S  Q  R  E  G  R  S
E10: GTCACGTTGTCTACCACTTACCTGAGACATTCTGATTTGGATGATGCTAGA
 +3:   H  V  V  Y  H  L  P  E  T  F  *  F  G  *  C  *  S
E10: GCACAGAAAATAGGTGAGCTGCATGGGATCCCAAAGCTGCTGAGGGATAGA
 +3:   T  E  N  R  *  A  A  W  D  P  K  A  A  E  G  *  S
E10: GCCTGAGCCTGGTGGCCACAGCATATGCCCTTTCTGTTCCATGCAGCTGGG
 +3:   L  S  L  V  A  T  A  Y  A  L  S  V  P  C  S  W  G
E10: GCTGTTAGTAGTCATTGCCCTTGTCAGCAGACCTTCTACCCTGGTGGCAAA
 +3:   C  *  *  S  L  P  L  S  A  D  L  L  P  W  W  Q  T
E10: CACATGAAAGCTGTGGCCCTGGGAGTGGCCTCCTAAAACAAGCCACTTAGG
 +3:   H  E  S  C  G  P  G  S  G  L  L  K  Q  A  T  *  V
E10: TCATCTGCCATCTACCCTTAACCTCTGTCTCTCGCCTGAGGGGAATCTGCA
 +3:   I  C  H  L  P  L  T  S  V  S  R  L  R  G  I  C  K
E10: AGCTGTGCATTGGGCTTACCTCCTGCTTTTGTAGAAATAACCATCCTTTGG
 +3:   L  C  I  G  L  T  S  C  F  C  R  N  N  H  P  L  V
E10: TATACATGGAGGATAGTTCCAGAACGCCTGAGTATACAAAAACCCAATGCA
 +3:   Y  M  E  D  S  S  R  T  P  E  Y  T  K  T  Q  C  I
E10: TACTCAAGTCCCACAGTGGGCCCTACAGAACCCACGTATGTGATAAATCAG
 +3:   L  K  S  H  S  G  P  Y  R  T  H  V  C  D  K  S  A
E10: CCCTCCATGTACGCAGGTTTCGCCCCCTGCCAATACTGTATTTTCAACCTG
 +3:   L  H  V  R  R  F  R  P  L  P  I  L  Y  F  Q  P  V
E10: TATGGTTGAAAAAAATCCATATATAAGTGCAGCCATGCAGTTCAAACCCAT
 +3:   W  L  K  K  I  H  I  *  V  Q  P  C  S  S  N  P  Y
E10: ATTGTTCAAGGGTCAACTGTATAGTTTATTGAACAGCCACACCCATTCCTT
 +3:   C  S  R  V  N  C  I  V  Y  *  T  A  T  P  I  P  L
E10: TACACATGATCTATGGCAGAGTTGAATAGTTGCAACAGACACTATGTGGCC
 +3:   H  M  I  Y  G  R  V  E  *  L  Q  Q  T  L  C  G  L
```

Figure 5E

```
E10: TGCAAAATCGGAAATTTTTACTGTCTGGCCTTTTACAGAAAAGTTTGCCAG
 +3:   Q  N  R  K  F  L  L  S  G  L  L  Q  K  S  L  P  A
E10: CCCCTGATCTAGACCAGCAGCTCATCTGATAGAGGCAGAGGTGGCCTTAAA
 +3:   P  D  L  D  Q  Q  L  I  *  *  R  Q  R  W  P  *  R
E10: GATGTGGCCTTCTTCATTTTCTGTTGGTTTGGTTTCGTTTCTATGAGAGAT
 +3:   C  G  L  L  H  F  L  L  V  W  F  R  F  Y  E  R  F
E10: TTCCTCTGATAGCTCTGCTTTCCCCAGCACTTACTCTCTGAGCTTTTAAAT
 +3:   P  L  I  A  L  L  S  P  A  L  T  L  *  A  F  K  C
E10: GTTCTCTCTGGGAGCTTCATATAAGCTCGGTGACATTTGAGCCACAGTTTT
 +3:   S  L  W  E  L  H  I  S  S  V  T  F  E  P  Q  F  L
E10: TAGATCAGCACCTGGAATACATGACACATTCTTACTGAGGTCATCCAGCAC
 +3:   D  Q  H  L  E  Y  M  T  H  S  Y  *  G  H  P  A  L
E10: TGCCATGGTGGCTGCCCAGTCTTCTGGCCAGTGTGCCAGGCACATGTCCCT
 +3:   P  W  W  L  P  S  L  L  A  S  V  P  G  T  C  P  C
E10: GTCACACAGGTTCCAAGAAACACATACGCAGCCATGCATAGACCAACAGAT
 +3:   H  T  G  S  K  K  H  I  R  S  H  A  *  T  N  R  F
E10: TTAATATTATATTGCAGTTTTCAGCGATGCAGAATGCAGCTGCAATTGTGT
 +3:   N  I  I  L  Q  F  S  A  M  Q  N  A  A  A  I  V  F
E10: TTTAAGGAGAAGCCAAATGGGGATGGTTGTCCCTGCAACATGGTGCCACTC
 +3:   *  G  E  A  K  W  G  W  L  S  L  Q  H  G  A  T  P
E10: CTGGGCCATGTGCAGCCTCAGTGGACACTCTTCCATAGCGCTGAGGCCCTG
 +3:   G  P  C  A  A  S  V  D  T  L  P  *  R  *  G  P  G
E10: GCCCCGCCTCCAGTTACCCTGTACTGCCCACTGCCTTACAGTTCAGTGCGC
 +3:   P  A  S  S  Y  P  V  L  P  T  A  L  Q  F  S  A  Q
E10: AGGCCTTCACCTTTTCATCACCAGCCTCTCTGCTCAGTGCTCTGGAGTTCT
 +3:   A  F  T  F  S  S  P  A  S  L  L  S  A  L  E  F  L
E10: TGACCTTGTCCTTTATCATGAGATTTGCTGAAATCACTAATGAAAATAACT
 +3:   T  L  S  F  I  M  R  F  A  E  I  T  N  E  N  N  S
E10: CCCAAAAGCAACAAACAAAAATATTAGTTTAACTGGCACTGTGGTATATTA
 +3:   Q  K  Q  Q  T  K  I  L  V  *  L  A  L  W  Y  I  K
E10: AAAGGCACAAGGGCATTGTGGCTTAACACTTTTGCTGGATCCCAAGAGACG
 +3:   R  H  K  G  I  V  A  *  H  F  C  W  I  P  R  D  A
E10: CACATGATGTTAAAAAGAGATCTGGCAGCAGTACTAATACTACATTTCAGT
 +3:   H  D  V  K  K  R  S  G  S  S  T  N  T  T  F  Q  C
E10: GTAATCATCTTGGGGTGGTTTGGCCAGGATTTCCCAATTCCTTGATATCTG
 +3:   N  H  L  G  V  V  W  P  G  F  P  N  S  L  I  S  G
E10: GAGTTTCTTCACCATTGTCCGGCATCCTGCGGAGGCTTAATATACAGGCGT
 +3:   V  S  S  P  L  S  G  I  L  R  R  L  N  I  Q  A  *
E10: AAGGTCAGCAGCAATTTGTCTAATAAGTGATGAGATCAGTAGCTGAAGTCT
 +3:   G  Q  Q  Q  F  V  *  *  V  M  R  S  V  A  E  V  S
E10: CTAAGCTGGGCCATTACTAAATACCATAGCCATGTTGATCTGGAAATTTAT
 +3:   K  L  G  H  Y  *  I  P  *  P  C  *  S  G  N  L  S
E10: CCCTCTAGTGTCTTACCTCACATAAGCCATTTGCCCACTGTGCAATATAGA
 +3:   L  *  C  L  T  S  H  K  P  F  A  H  C  A  I  *  K
E10: AAGGTGTTTTCAAAAGTATTTGGCCGTAGATTTTCACATCCATCATAAGGT
 +3:   G  V  F  K  S  I  W  P  *  I  F  T  S  I  I  R  L
```

THIS PAGE BLANK (USPTO)

Figure 5F

```
E10: TGGCATTCAATAAGGAAAAAGTTCTAACTCCAGTATTAAATTGTACATAAA
 +3:   A  F  N  K  E  K  V  L  T  P  V  L  N  C  T  *  I
E10: TCCCAAATGTTCTTAAAGAACACTCAGGGACATGTTTGTTGCCTGGGATTG
 +3:   P  N  V  L  K  E  H  S  G  T  C  L  L  P  G  I  G
E10: GTAATGAAAGGTTGGTTTTTGAAACTTGAAATTTCACCATTGGTTTTTTTC
 +3:   N  E  R  L  V  F  E  T  *  N  F  T  I  G  F  F  P
E10: CTATCATTTCTGCATATCCAGCAAAAGGAATCTCATGTTGACTCCTGGCAG
 +3:   I  I  S  A  Y  P  A  K  G  I  S  C  *  L  L  A  E
E10: AGTTCAGTGGCTTCAGTCTGTCTATCTGTTCTGAGGGGAAAATTGTGTTCT
 +3:   F  S  G  F  S  L  S  I  C  S  E  G  K  I  V  F  W
E10: GGATCCAGTAATCAATTTGGCAACTTTAATCGAGGTTTTCAAAATTCCAAG
 +3:   I  Q  *  S  I  W  Q  L  *  S  R  F  S  K  F  Q  G
E10: GAGGGTTAATAAAGAATGATAATCAGTTTTATTTGCTAATAGCTAAGACAA
 +3:   G  L  I  K  N  D  N  Q  F  Y  L  L  I  A  K  T  N
E10: ATTTGTAATAAAGTGTTTTATAATACTTC
 +3:   L  *  *  S  V  L  *  Y  F
```

...gtgcttttctctttag

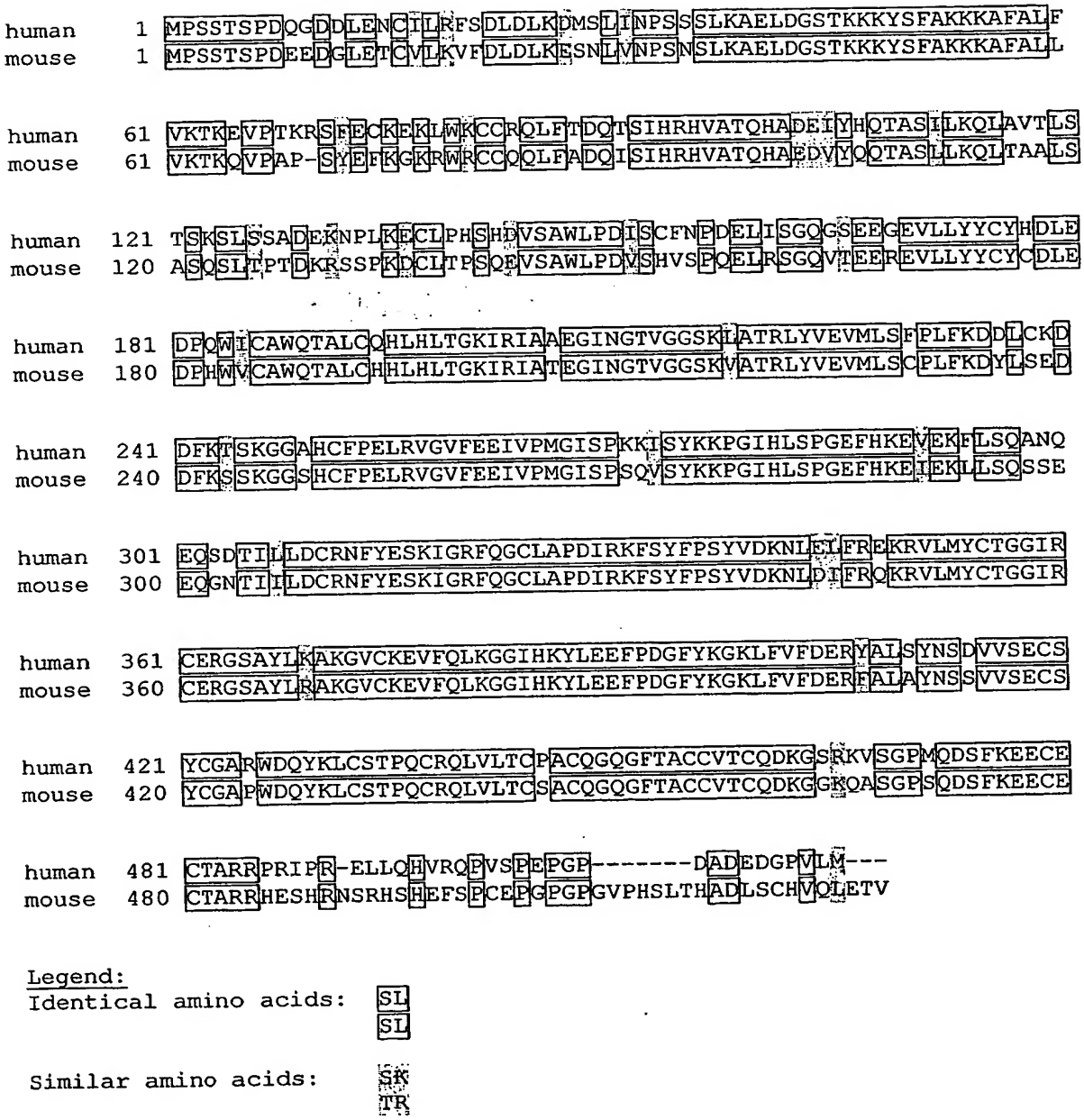
```
E6a: GTTACGAGACAGTACAATAGAAGGAGTATGCTCGTCCCCATTCTTTCACTG
 +1: V  T  R  Q  Y  N  R  R  S  M  L  V  P  I  L  S  L
E6a: AGTCACCATATGATTTTGGACCAGCTAGTGCTCTAGACCTCAGTATCCCTT
 +1: S  H  H  M  I  L  D  Q  L  V  L  *  T  S  V  S  L
E6a: CTTATAAAATAAGAATGTTACAGCTCATGCAATCTGGGACTCCAAATCTTG
 +1: L  I  K  *  E  C  Y  S  S  C  N  L  G  L  Q  I  L
E6a: GACATATTAGCTCACTTGAGAGACCACCAGCCTGGTCAGCAGATCACTGTG
 +1: D  I  L  A  H  L  R  D  H  Q  P  G  Q  Q  I  T  V
E6a: TTTTTAGTAAATCTGGAATTGTAAGATTAACACTTCATACCACATGGGGGA
 +1: F  L  V  N  L  E  L  *  D  *  H  F  I  P  H  G  G
E6a: ATAAAGTTGTTGCTCTCACAGGT
 +1: I  K  L  L  L  S  Q
```

gggctg...

Figure 5G

```
          ...gtttttccacacctag

E9a: GGAGTGTGCAAGGAGGTGTTCCAGCTCAAGGGTGGCATCCACAAGTACCTG
 +1: G  V  C  K  E  V  F  Q  L  K  G  G  I  H  K  Y  L
E9a: GAAGAGTTTCCTGATGGCTTTTACAAAGGGAAGTTGTTTGTTTTTGATGAA
 +1: E  E  F  P  D  G  F  Y  K  G  K  L  F  V  F  D  E
E9a: CGCTATGCTCTGTCCTACAACAGTGATGTGGTGTCAGGTAGGTCAGCACAG
 +1: R  Y  A  L  S  Y  N  S  D  V  V  S  G  R  S  A  Q
E9a: GCTCAGAGCCCAAACTGAAATGAAGCACATTGTCAGTTCACTATTCTAGAA
 +1: A  Q  S  P  N  *  N  E  A  H  C  Q  F  T  I  L  E
E9a: AAATGACACAGGGAAGACAGGCCAGTGCTCATTACTGAGCACTGAATAAGC
 +1: K  *  H  R  E  D  R  P  V  L  I  T  E  H  *  I  S
E9a: AGGGAAAATAAGTACATTGTGCCACCATTTTCCCAGCTGTGGAGCTGAGAG
 +1: R  E  N  K  Y  I  V  P  P  F  S  Q  L  W  S  *  E
E9a: AACCCTAGCCCAGGAGTCAGGAGGCCTGGGTTGGGATCCTGGCTTCACCAT
 +1: N  P  S  P  G  V  R  R  P  G  L  G  S  W  L  H  H
E9a: TGCTAGCTGGACAAGCCCATTAACATGGGGATCATCTCACCTGCCCTGCCT
 +1: C  *  L  D  K  P  I  N  M  G  I  I  S  P  A  L  P
E9a: GCCTGTCTACCTGCCAAGAGCTGTACTACTGGGCTAATTCAGGGCTCTTAA
 +1: A  C  L  P  A  K  S  C  T  T  G  L  I  Q  G  S  *
E9a: CCTGGAATTGGTACATAGATTTCAGGGATTCTGTGAATTTGGATGGAAAAA
 +1: P  G  I  G  T  *  I  S  G  I  L  *  I  W  M  E  K
E9a: TAATTGTATCTTTGTTTTCAATAACACCTCACTAAAATGAAGCATTTCCTT
 +1: *  L  Y  L  C  F  Q  *  H  L  T  K  M  K  H  F  L
E9a: TAGTTATGAATGTAGGCAACAAAGTACCAGTTGTATTAATGTACCTGTGAC
 +1: *  L  *  M  *  A  T  K  Y  Q  L  Y  *  C  T  C  D
E9a: TTTGTCTTCAGTAGGATTCACAATACTTTCATATCATGTTCTAGTTGCCTC
 +1: F  V  F  S  R  I  H  N  T  F  I  S  C  S  S  C  L
E9a: AGATATCTCAAAATAGTATTTATACTCATCACTGCTTCAAAATGAAAATAG
 +1: R  Y  L  K  I  V  F  I  L  I  T  A  S  K  *  K  *
E9a: TTATTAGGCCCACCACTAAGAGTTGATATATAATGTGTTAATAAATGGCAC
 +1: L  L  G  P  P  L  R  V  D  I  *  C  V  N  K  W  H
E9a: GTCTTATTATATATTACAGATTTTGAAAAAGA
 +1: V  L  L  Y  I  T  D  F  E  K

                    ...ctttga
```

Figure 6

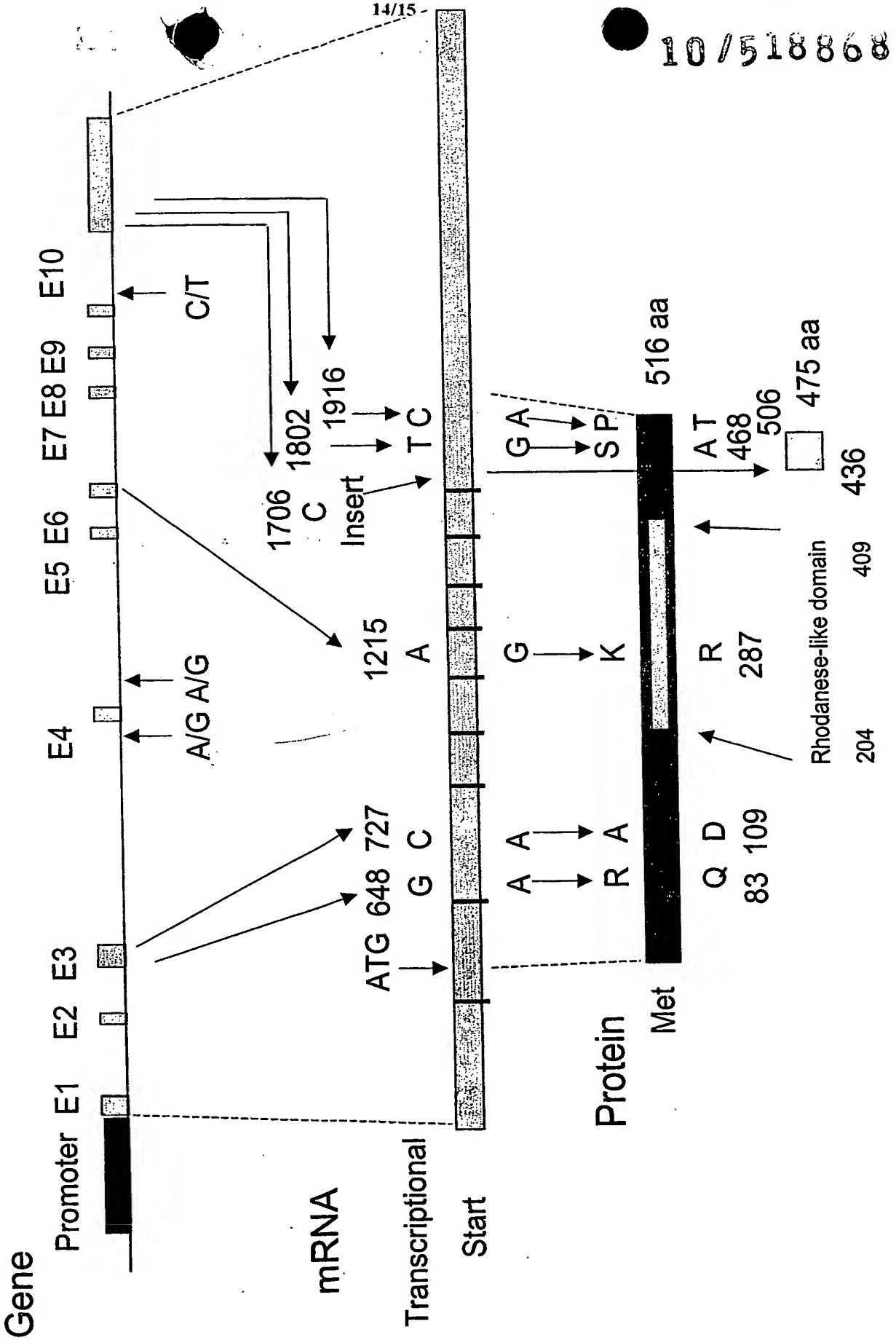


THIS PAGE BLANK (USPTO)

Figure 7

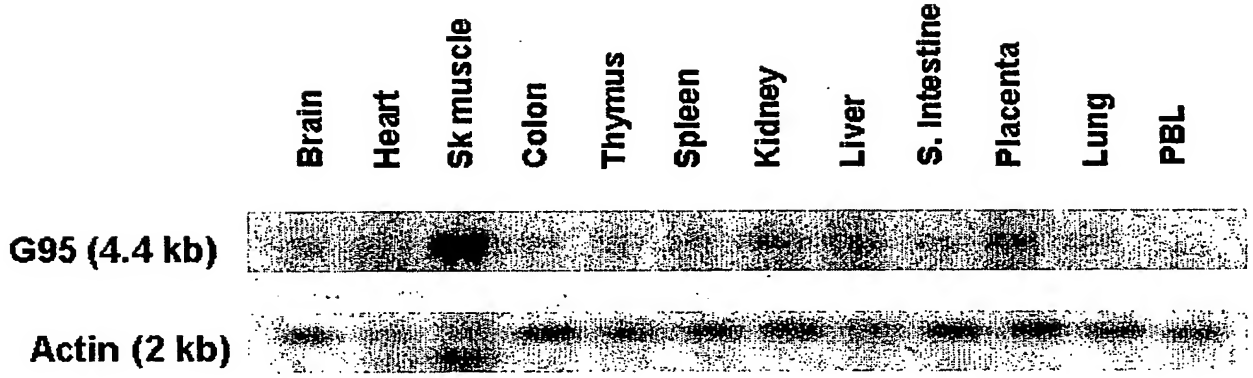
```
P1G95-1  LKGISASLLQPNQMPSSTSPDQGDDLENCILRFSDLDLKDMSLINPSSSLKAELDGSTKK
P1G95-2  LKGISASLLQPNQMPSSTSPDQGDDLENCILRFSDLDLKDMSLINPSSSLKAELDGSTKK
P1G95-n  LKGISASLLQPNQMPSSTSPDQGDDLENCILRFSDLDLKDMSLINPSSSLKAELDGSTKK
         ************************************************************

P1G95-1  KYSFAKKKAFALFVKTKEVPTKRSFECKEKLWKCCRQLFTDQTSIHRHVATQHADEIYHQ
P1G95-2  KYSFAKKKAFALFVKTKEVPTKRSFECKEKLWKCCRQLFTDQTSIHRHVATQHADEIYHQ
P1G95-n  KYSFAKKKAFALFVKTKEVPTKRSFECKEKLWKCCRQLFTDQTSIHRHVATQHADEIYHQ
         ************************************************************

P1G95-1  TASILKQLAVTLSTSKSLSSADEKNPLKECLPHSHDVSAWLPDISCFNPDELISGQGSEE
P1G95-2  TASILKQLAVTLSTSKSLSSADEKNPLKECLPHSHDVSAWLPDISCFNPDELISGQGSEE
P1G95-n  TASILKQLAVTLSTSKSLSSADEKNPLKECLPHSHDVSAWLPDISCFNPDELISGQGSEE
         ************************************************************

P1G95-1  GEVLLYYCYHDLEDPQWICAWQTALCQHLHLTGKIRIAAEGINGTVGGSKLATRLYVEVM
P1G95-2  GEVLLYYCYHDLEDPQWICAWQTALCQHLHLTGKIRIAAEGINGTVGGSKLATRLYVEVM
P1G95-n  GEVLLYYCYHDLEDPQWICAWQTALCQHLHLTGKIRIAAEGINGTVGGSKLATRLYVEVM
         ************************************************************

P1G95-1  LSFPLFKDDLCKDDFKTSKGGAHCFPELRVGVFEEIVPMGISPKKISYKKPGIHLSPGEF
P1G95-2  LSFPLFKDDLCKDDFKTSKGGAHCFPELRVGVFEEIVPMGISPKKISYKKPGIHLSPGEF
P1G95-n  LSFPLFKDDLCKDDFKTSKGGAHCFPELRVGVFEEIVPMGISPKKISYKKPGIHLSPGEF
         ************************************************************

P1G95-1  HKEVEKFLSQANQEQSDTILLDCRNFYESKIGRFQGCLAPDIRKFSYFPSYVDKNLELFR
P1G95-2  HKEVEKFLSQANQEQSDTILLDCRNFYESKIGRFQGCLAPDIRKFSYFPSYVDKNLELFR
P1G95-n  HKEVEKFLSQANQEQSDTILLDCRNFYESKIGRFQGCLAPDIRKFSYFPSYVDKNLELFR
         ************************************************************

P1G95-1  EKRVLMYCTGGIRCERGSAYLKAKGVCKEVFQLKGGIHKYLEEFPDGFYKGKLFVFDERY
P1G95-2  EKRVLMYCTGGIRCERGSAYLKAKGVCKEVFQLKGGIHKYLEEFPDGFYKGKLFVFDERY
P1G95-n  EKRVLMYCTGGIRCERGSAYLKAKGVCKEVFQLKGGIHKYLEEFPDGFYKGKLFVFDERY
         ************************************************************

P1G95-1  ALSYNSDVVSECSYCGARWDQYKLCSTPQCRQLVLTCPACQGQGFTACCVTCQDKGSRKV
P1G95-2  ALSYNSDVVSECSYCGARWDQYKLCSTPPVPPARFDLPCLSRTRIHSLLCHMSRQGEQES
P1G95-n  ALSYNSDVVSECSYCGARWDQYKLCSTPQCRQLVLTCPACQGQGFTACCVTCQDKGSRKV
         ****************************      :  *. .   : :     . :*.::

P1G95-1  AGPMQDSFKEECECTARRPRIPRELLQHVRQPVSPEPGPDADEDGPVLM
P1G95-2  FRPYARQL----------------------------------------
P1G95-n  SGPMQDSFKEECECTARRPRIPRELLQHVRQPVSPEPGPDADEDGPVLM
           *   .:   . .:   .  .          ..:. ... *   ...
```

Figure 8


Gene
Promoter
E1
E2
E3
E4
E5
E6
E7
E8
E9
E10
C/T
A/G
A/G
1706
C
1802
1916
Insert
T
C
1215
A
G
K
ATG
648
727
G
C
A
A
R
A
Q
D
83
109
Transcriptional
Start
mRNA
Protein
Met
G
A
S
P
516 aa
A
T
468
506
475 aa
436
R
287
Rhodanese-like domain
409
204

THIS PAGE BLANK (USPTO)

FIGURE 9



THIS PAGE BLANK (USPTO)

SEQUENCE LISTING

<110> Xenon Genetics, Inc.
Warner-Lambert Company, LLC

<120> Novel Therapeutic Target for Treating Vascular Diseases, Dyslipidemias and Related Disorders

<130> 760050-100

<140>
<141>

<150> US/60/391,878
<151> 2002-06-27

<160> 32

<170> PatentIn version 3.0

<210> 1
<211> 501
<212> DNA
<213> Homo sapiens

<400> 1
agtagctctc ctgataaaag accaccaacc atgggctagg tctggccagt ttacagaaag 60
cacacactgt gtgcctttat gtcctagaaa gatcttttga tatacaggac ctaaatggaa 120
tacattccac cccaaaataa acatgggtca tacatgcata tttattcaat acacatatgt 180
caggaccatc ttcataaata ttcatagctc ctcctataat ctgttaaata tgtgtgtgtg 240
tgcgtgtgtg tgtgtgtgtg tgtgtgtgtg tatagtttgt ttgttttgag agggagtctt 300
gctctgttgc ccaggctgga gtgcagcagt gcaatctcaa ctcactataa cttccacctc 360
caggttcaag caattctcat gtctcagccg agtagctggg accacagtca catgccacca 420
ctcttggcta attttttttt ttttttttttg agacggagtc tcgctctgtc acccaggcta 480
gagtgcagtg gcgcgatctt g 501

<210> 2
<211> 600
<212> DNA
<213> Homo sapiens

<400> 2
gctgctgcct ggatgaactt caagtgttcc ttctcctgtt ccctgctaca tcttagattt 60
tacagtgtgt ccttggctaa aggcagcctg ggcacactgc agtgccctgg tgtaacggac 120
agccatgggc cttgcacttg aactagggtc tggccccagg actgtgacat gctcaccctg 180
agccccgggt tctcctctga gaaatgctgg ggtcacctgc tttgagggct gttcttagta 240
tgaagcaaga gcacagtaag acaaaagact acagagccga cacacacaca cacacacaca 300
cacacacaca cacacacaca cagagtagtg cattccagaa caatatactc atttcatttt 360
cctgttgtca ttcagagagg cgagtgcact gggagccaca aaagtgcaat gttgcaaaga 420
cttttccaaa acaggatgcg taatggctgc tgtggccact gctggcggtg tgtggggata 480
cctgggtagc agcaggccac cagagagtgt gcatccctcc ttctgtgctc tgcagtgggg 540
ctcatttttc tgggccaggt tcttgcctgc ctgcaatcct cctctgatga cggagttacc 600

<210> 3
<211> 4333
<212> DNA
<213> Homo sapiens

<400> 3

atgtccttcc ggggccacgc agaaagtgcc gccgctttgg ccactcagag cccccgggcc 60
gcggtcgtcg tacgcctgaa ggcgggtcgt gccggcggcc gctctagtct ccgcctccgc 120
tcaggccggt cctccggggc ttctcaatgg tttcccggtg gcctctcaat ggttttcccg 180
gcggcccttg cgccgacgcc aggagacttc cggagcttgg tgacgtcacg agcgagcttt 240
tctacccaaa tacgcggcgg gggaataggc tcgagggcgg tgagcagtga caattgctag 300
gcggagacag tgcagggaag agagacctta gaaaggatca ggactggcgg gaggtattta 360
actgaaagga atatctgctt cactgttgca accaaaccag atgccttctt ccacttcacc 420
agaccaagga gatgacctgg agaactgcat tttaagattt tctgacctgg atttaaaaga 480
tatgagtctt attaatccca gcagcagtct taaagcagaa ttagatggca gtacaaaaaa 540
gaaatactcg tttgcaaaga aaaaggcctt tgccctttt gtcaaaacca aagaagttcc 600
aacaaaaagg agttttgaat gtaaagaaaa attgtggaaa tgctgtcggc agctattcac 660
agaccaaacc agcatccata gacatgtggc aacacaacat gctgatgaaa tttatcacca 720
gacagcttct attttaaagc aactggctgt gacattgagc acctcaaaga gtctttcgtc 780
tgcagatgaa aagaaccctt taaaagagtg ccttccacat agccatgacg tgtctgcttg 840
gctccctgat ataagctgct ttaaccctga tgagctgata agtggccagg gcagtgaaga 900
aggggaggtg ctcctttatt actgctacca tgacctggag gatccccaat ggatctgtgc 960
ctggcagaca gctctgtgtc agcacctgca cctcacaggc aagattcgaa ttgctgcaga 1020
aggaatcaat gggacagttg gtggaagcaa attggctacc agactttatg tggaagtcat 1080
gctttccttc ccattgttta aggatgacct gtgtaaagat gattttaaga ccagcaaagg 1140
aggagctcac tgttttccag aattgcgtgt tggtgtattt gaagaaatcg tgcccatggg 1200
gatcagcccc aaaaagatct cctacaagaa gcctggaatc catttatccc caggtgaatt 1260
tcataaagaa gtagaaaagt ttttatctca ggcaaatcaa gaacaaagtg atactatcct 1320
tcttgattgc agaaacttct atgaaagcaa aataggacga ttccaaggct gcttagcccc 1380
agacatcagg aaattcagtt acttccctag ctacgttgac aaaaatctag aacttttcag 1440
agagaagaga gtgctgatgt actgtaccgg gggcatccgc tgtgagcggg gttcagccta 1500
cctcaaagcc aagggagtgt gcaaggaggt gttccagctc aagggtggca tccacaagta 1560
cctggaagag tttcctgatg gcttttacaa agggaagttg tttgtttttg atgaacgcta 1620
tgctctgtcc tacaacagtg atgtggtgtc agagtgttca tactgtggag cccgctggga 1680
ccagtataaa ctctgctcta ctccccagtg ccgccagctc gttttgacct gccctgcctg 1740
tcaaggacaa ggattcacag cctgttgtgt cacatgtcaa gacaagggga gcaggaaagt 1800
ttcaggccct atgcaagaca gctttaaaga ggaatgcgag tgcacagccc gacggccacg 1860
catacctagg gaactcttgc agcatgtgcg acagcctgtg agcccagagc cagggcctga 1920
tgctgatgag gatgggccag tgcttatgtg agcagcacct ttggcatttt cccaggccct 1980
cggtaaaagt aggtttgggg tgactataca gagaaagcat ggcaagactg cagaaacaga 2040
gaaatcggga acttcagttc tggccgctgc caccgtggca gccgtctaca cttcacagcg 2100
ggaggggagg agtcacgttg tctaccactt acctgagaca ttctgatttg gatgatgcta 2160
gagcacagaa aataggtgag ctgcatggga tcccaaagct gctgagggat agagcctgag 2220
cctggtggcc acagcatatg ccctttctgt tccatgcagc tggggctgtt agtagtcatt 2280
gcccttgtca gcagaccttc taccctggtg gcaaacacat gaaagctgtg gccctgggag 2340
tggcctccta aaacaagcca cttaggtcat ctgccatcta cccttaacct ctgtctctcg 2400
cctgagggga atctgcaagc tgtgcattgg gcttacctcc tgcttttgta gaaataacca 2460
tcctttggta tacatggagg atagttccag aacgcctgag tatacaaaaa cccaatgcat 2520
actcaagtcc cacagtgggc cctacagaac ccacgtatgt gataaatcag ccctccatgt 2580
acgcaggttt cgccccctgc caatactgta ttttcaacct gtatggttga aaaaaatcca 2640
tatataagtg cagccatgca gttcaaaccc atattgttca agggtcaact gtatagttta 2700
ttgaacagcc acacccattc ctttacacat gatctatggc agagttgaat agttgcaaca 2760
gacactatgt ggcctgcaaa atcggaaatt tttactgtct ggccttttac agaaaagttt 2820
gccagcccct gatctagacc agcagctcat ctgatagagg cagaggtggc cttaaagatg 2880
tggccttctt cattttctgt tggtttggtt tcgtttctat gagagatttc ctctgatagc 2940
tctgctttcc ccagcactta ctctctgagc ttttaaatgt tctctctggg agcttcatat 3000
aagctcggtg acatttgagc cacagttttt agatcagcac ctggaataca tgacacattc 3060
ttactgaggt catccagcac tgccatggtg gctgcccagt cttctggcca gtgtgccagg 3120
cacatgtccc tgtcacacag gttccaagaa acacatacgc agccatgcat agaccaacag 3180
atttaatatt atattgcagt tttcagcgat gcagaatgca gctgcaattg tgttttaagg 3240
agaagccaaa tggggatggt tgtccctgca acatggtgcc actcctgggc catgtgcagc 3300
ctcagtggac actcttccat agcgctgagg ccctggcccc gcctccagtt accctgtact 3360
gcccactgcc ttacagttca gtgcgcaggc cttcaccttt tcatcaccag cctctctgct 3420
cagtgctctg gagttcttga ccttgtcctt tatcatgaga tttgctgaaa tcactaatga 3480
aaataactcc caaaagcaac aaacaaaaat attagtttaa ctggcactgt ggtatattaa 3540
aaggcacaag ggcattgtgg cttaacactt ttgctggatc ccaagagacg cacatgatgt 3600

taaaaagaga tctggcagca gtactaatac tacatttcag tgtaatcatc ttggggtggt 3660
ttggccagga tttcccaatt ccttgatatc tggagtttct tcaccattgt ccggcatcct 3720
gcggaggctt aatatacagg cgtaaggtca gcagcaattt gtctaataag tgatgagatc 3780
agtagctgaa gtctctaagc tgggccatta ctaaatacca tagccatgtt gatctggaaa 3840
tttatccctc tagtgtctta cctcacataa gccatttgcc cactgtgcaa tatagaaagg 3900
tgttttcaaa agtatttggc cgtagatttt cacatccatc ataaggttgg cattcaataa 3960
ggaaaaagtt ctaactccag tattaaattg tacataaatc ccaaatgttc ttaaagaaca 4020
ctcagggaca tgtttgttgc ctgggattgg taatgaaagg ttggtttttg aaacttgaaa 4080
tttcaccatt ggtttttttc ctatcatttc tgcatatcca gcaaaaggaa tctcatgttg 4140
actcctggca gagttcagtg gcttcagtct gtctatctgt tctgagggga aaattgtgtt 4200
ctggatccag taatcaattt ggcaacttta atcgaggttt tcaaaattcc aaggagggtt 4260
aataaagaat gataatcagt tttatttgct aatagctaag acaaatttgt aataaagtgt 4320
tttataatac ttc 4333

<210> 4
<211> 516
<212> PRT
<213> Homo sapiens

<400> 4
Met Pro Ser Ser Thr Ser Pro Asp Gln Gly Asp Asp Leu Glu Asn Cys
1 5 10 15

Ile Leu Arg Phe Ser Asp Leu Asp Leu Lys Asp Met Ser Leu Ile Asn
20 25 30

Pro Ser Ser Ser Leu Lys Ala Glu Leu Asp Gly Ser Thr Lys Lys Lys
35 40 45

Tyr Ser Phe Ala Lys Lys Lys Ala Phe Ala Leu Phe Val Lys Thr Lys
50 55 60

Glu Val Pro Thr Lys Arg Ser Phe Glu Cys Lys Glu Lys Leu Trp Lys
65 70 75 80

Cys Cys Arg Gln Leu Phe Thr Asp Gln Thr Ser Ile His Arg His Val
85 90 95

Ala Thr Gln His Ala Asp Glu Ile Tyr His Gln Thr Ala Ser Ile Leu
100 105 110

Lys Gln Leu Ala Val Thr Leu Ser Thr Ser Lys Ser Leu Ser Ser Ala
115 120 125

Asp Glu Lys Asn Pro Leu Lys Glu Cys Leu Pro His Ser His Asp Val
130 135 140

Ser Ala Trp Leu Pro Asp Ile Ser Cys Phe Asn Pro Asp Glu Leu Ile
145 150 155 160

Ser Gly Gln Gly Ser Glu Glu Gly Glu Val Leu Leu Tyr Tyr Cys Tyr
165 170 175

His Asp Leu Glu Asp Pro Gln Trp Ile Cys Ala Trp Gln Thr Ala Leu
180 185 190

Cys Gln His Leu His Leu Thr Gly Lys Ile Arg Ile Ala Ala Glu Gly
195 200 205

Ile Asn Gly Thr Val Gly Gly Ser Lys Leu Ala Thr Arg Leu Tyr Val

210 215 220

Glu Val Met Leu Ser Phe Pro Leu Phe Lys Asp Asp Leu Cys Lys Asp
225 230 235 240

Asp Phe Lys Thr Ser Lys Gly Gly Ala His Cys Phe Pro Glu Leu Arg
245 250 255

Val Gly Val Phe Glu Glu Ile Val Pro Met Gly Ile Ser Pro Lys Lys
260 265 270

Ile Ser Tyr Lys Lys Pro Gly Ile His Leu Ser Pro Gly Glu Phe His
275 280 285

Lys Glu Val Glu Lys Phe Leu Ser Gln Ala Asn Gln Glu Gln Ser Asp
290 295 300

Thr Ile Leu Leu Asp Cys Arg Asn Phe Tyr Glu Ser Lys Ile Gly Arg
305 310 315 320

Phe Gln Gly Cys Leu Ala Pro Asp Ile Arg Lys Phe Ser Tyr Phe Pro
325 330 335

Ser Tyr Val Asp Lys Asn Leu Glu Leu Phe Arg Glu Lys Arg Val Leu
340 345 350

Met Tyr Cys Thr Gly Gly Ile Arg Cys Glu Arg Gly Ser Ala Tyr Leu
355 360 365

Lys Ala Lys Gly Val Cys Lys Glu Val Phe Gln Leu Lys Gly Gly Ile
370 375 380

His Lys Tyr Leu Glu Glu Phe Pro Asp Gly Phe Tyr Lys Gly Lys Leu
385 390 395 400

Phe Val Phe Asp Glu Arg Tyr Ala Leu Ser Tyr Asn Ser Asp Val Val
405 410 415

Ser Glu Cys Ser Tyr Cys Gly Ala Arg Trp Asp Gln Tyr Lys Leu Cys
420 425 430

Ser Thr Pro Gln Cys Arg Gln Leu Val Leu Thr Cys Pro Ala Cys Gln
435 440 445

Gly Gln Gly Phe Thr Ala Cys Cys Val Thr Cys Gln Asp Lys Gly Ser
450 455 460

Arg Lys Val Ser Gly Pro Met Gln Asp Ser Phe Lys Glu Glu Cys Glu
465 470 475 480

Cys Thr Ala Arg Arg Pro Arg Ile Pro Arg Glu Leu Leu Gln His Val
485 490 495

Arg Gln Pro Val Ser Pro Glu Pro Gly Pro Asp Ala Asp Glu Asp Gly
500 505 510

Pro Val Leu Met
515

<210> 5

<211> 526
<212> PRT
<213> Mus musculus

<400> 5

```
Met Pro Ser Ser Thr Ser Pro Asp Glu Glu Asp Gly Leu Glu Thr Cys
1               5                  10                  15

Val Leu Lys Val Phe Asp Leu Asp Leu Lys Glu Ser Asn Leu Val Asn
            20                  25                  30

Pro Ser Asn Ser Leu Lys Ala Glu Leu Asp Gly Ser Thr Lys Lys Lys
        35                  40                  45

Tyr Ser Phe Ala Lys Lys Lys Ala Phe Ala Leu Leu Val Lys Thr Lys
    50                  55                  60

Gln Val Pro Ala Pro Ser Tyr Glu Phe Lys Gly Lys Arg Trp Arg Cys
65                  70                  75                  80

Cys Gln Gln Leu Phe Ala Asp Gln Ile Ser Ile His Arg His Val Ala
                85                  90                  95

Thr Gln His Ala Glu Asp Val Tyr Gln Gln Thr Ala Ser Leu Leu Lys
            100                 105                 110

Gln Leu Thr Ala Ala Leu Ser Ala Ser Gln Ser Leu Thr Pro Thr Asp
        115                 120                 125

Lys Arg Ser Ser Pro Lys Asp Cys Leu Thr Pro Ser Gln Glu Val Ser
    130                 135                 140

Ala Trp Leu Pro Asp Val Ser His Val Ser Pro Gln Glu Leu Arg Ser
145                 150                 155                 160

Gly Gln Val Thr Glu Glu Arg Glu Val Leu Leu Tyr Tyr Cys Tyr Cys
                165                 170                 175

Asp Leu Glu Asp Pro His Trp Val Cys Ala Trp Gln Thr Ala Leu Cys
            180                 185                 190

His His Leu His Leu Thr Gly Lys Ile Arg Ile Ala Thr Glu Gly Ile
        195                 200                 205

Asn Gly Thr Val Gly Gly Ser Lys Val Ala Thr Arg Leu Tyr Val Glu
    210                 215                 220

Val Met Leu Ser Cys Pro Leu Phe Lys Asp Tyr Leu Ser Glu Asp Asp
225                 230                 235                 240

Phe Lys Ser Ser Lys Gly Gly Ser His Cys Phe Pro Glu Leu Arg Val
                245                 250                 255

Gly Val Phe Glu Glu Ile Val Pro Met Gly Ile Ser Pro Ser Gln Val
            260                 265                 270

Ser Tyr Lys Lys Pro Gly Ile His Leu Ser Pro Gly Glu Phe His Lys
        275                 280                 285

Glu Ile Glu Lys Leu Leu Ser Gln Ser Ser Glu Glu Gln Gly Asn Thr
    290                 295                 300
```

Ile Ile Leu Asp Cys Arg Asn Phe Tyr Glu Ser Lys Ile Gly Arg Phe
305 310 315 320

Gln Gly Cys Leu Ala Pro Asp Ile Arg Lys Phe Ser Tyr Phe Pro Ser
325 330 335

Tyr Val Asp Lys Asn Leu Asp Ile Phe Arg Gln Lys Arg Val Leu Met
340 345 350

Tyr Cys Thr Gly Gly Ile Arg Cys Glu Arg Gly Ser Ala Tyr Leu Arg
355 360 365

Ala Lys Gly Val Cys Lys Glu Val Phe Gln Leu Lys Gly Gly Ile His
370 375 380

Lys Tyr Leu Glu Glu Phe Pro Asp Gly Phe Tyr Lys Gly Lys Leu Phe
385 390 395 400

Val Phe Asp Glu Arg Phe Ala Leu Ala Tyr Asn Ser Ser Val Val Ser
405 410 415

Glu Cys Ser Tyr Cys Gly Ala Pro Trp Asp Gln Tyr Lys Leu Cys Ser
420 425 430

Thr Pro Gln Cys Arg Gln Leu Val Leu Thr Cys Ser Ala Cys Gln Gly
435 440 445

Gln Gly Phe Thr Ala Cys Cys Val Thr Cys Gln Asp Lys Gly Gly Lys
450 455 460

Gln Ala Ser Gly Pro Ser Gln Asp Ser Phe Lys Glu Glu Cys Glu Cys
465 470 475 480

Thr Ala Arg Arg His Glu Ser His Arg Asn Ser Arg His Ser His Glu
485 490 495

Phe Ser Pro Cys Glu Pro Gly Pro Gly Pro Gly Val Pro His Ser Leu
500 505 510

Thr His Ala Asp Leu Ser Cys His Val Gln Leu Glu Thr Val
515 520 525

<210> 6
<211> 4333
<212> DNA
<213> Homo sapiens

<400> 6
atgtccttcc ggggccacgc agaaagtgcc gccgctttgg ccactcagag cccccgggcc 60
gcggtcgtcg tacgcctgaa ggcgggtcgt gccggcggcc gctctagtct ccgcctccgc 120
tcaggccggt cctccggggc ttctcaatgg tttcccggtg gcctctcaat ggttttcccg 180
gcggcccttg cgccgacgcc aggagacttc cggagcttgg tgacgtcacg agcgagcttt 240
tctacccaaa tacgcggcgg gggaataggc tcgagggcgg tgagcagtga caattgctag 300
gcggagacag tgcagggaag agagacctta gaaaggatca ggactggcgg gaggtattta 360
actgaaagga atatctgctt cactgttgca accaaaccag atgccttctt ccacttcacc 420
agaccaagga gatgacctgg agaactgcat tttaagattt tctgacctgg atttaaaaga 480
tatgagtctt attaatccca gcagcagtct taaagcagaa ttagatggca gtacaaaaaa 540
gaaatactcg tttgcaaaga aaaaggcctt tgccctttt gtcaaaacca aagaagttcc 600

aacaaaaagg agtttgaat gtaaagaaaa attgtggaaa tgctgtcggc agctattcac 660
agaccaaacc agcatccata gacatgtggc aacacaacat gctgatgaaa tttatcacca 720
gacagcttct attttaaagc aactggctgt gacattgagc acctcaaaga gtctttcgtc 780
tgcagatgaa aagaaccctt taaaagagtg ccttccacat agccatgacg tgtctgcttg 840
gctccctgat ataagctgct ttaaccctga tgagctgata agtggccagg gcagtgaaga 900
aggggaggtg ctcctttatt actgctacca tgacctggag gatccccaat ggatctgtgc 960
ctggcagaca gctctgtgtc agcacctgca cctcacaggc aagattcgaa ttgctgcaga 1020
aggaatcaat gggacagttg gtggaagcaa attggctacc agactttatg tggaagtcat 1080
gctttccttc ccattgttta aggatgacct gtgtaaagat gattttaaga ccagcaaagg 1140
aggagctcac tgttttccag aattgcgtgt tggtgtattt gaagaaatcg tgcccatggg 1200
gatcagcccc aaaaagatct cctacaagaa gcctggaatc catttatccc caggtgaatt 1260
tcataaagaa gtagaaaagt ttttatctca ggcaaatcaa gaacaaagtg atactatcct 1320
tcttgattgc agaaacttct atgaaagcaa aataggacga ttccaaggct gcttagcccc 1380
agacatcagg aaattcagtt acttccctag ctacgttgac aaaaatctag aacttttcag 1440
agagaagaga gtgctgatgt actgtaccgg gggcatccgc tgtgagcggg gttcagccta 1500
cctcaaagcc aagggagtgt gcaaggaggt gttccagctc aagggtggca tccacaagta 1560
cctggaagag tttcctgatg gcttttacaa agggaagttg tttgtttttg atgaacgcta 1620
tgctctgtcc tacaacagtg atgtggtgtc agagtgttca tactgtggag cccgctggga 1680
ccagtataaa ctctgctcta ctccccagtg ccgccagctc gttttgacct gccctgcctg 1740
tcaaggacaa ggattcacag cctgttgtgt cacatgtcaa gacaagggga gcaggaaagt 1800
tgcaggccct atgcaagaca gctttaaaga ggaatgcgag tgcacagccc gacggccacg 1860
catacctagg gaactcttgc agcatgtgcg acagcctgtg agcccagagc cagggcctga 1920
tgctgatgag gatgggccag tgcttatgtg agcagcacct ttggcatttt cccaggccct 1980
cggtaaaagt aggtttgggg tgactataca gagaaagcat ggcaagactg cagaaacaga 2040
gaaatcggga acttcagttc tggccgctgc caccgtggca gccgtctaca cttcacagcg 2100
ggaggggagg agtcacgttg tctaccactt acctgagaca ttctgatttg gatgatgcta 2160
gagcacagaa aataggtgag ctgcatggga tcccaaagct gctgagggat agagcctgag 2220
cctggtggcc acagcatatg ccctttctgt tccatgcagc tggggctgtt agtagtcatt 2280
gcccttgtca gcagaccttc taccctggtg gcaaacacat gaaagctgtg gccctgggag 2340
tggcctccta aaacaagcca cttaggtcat ctgccatcta cccttaacct ctgtctctcg 2400
cctgagggga atctgcaagc tgtgcattgg gcttacctcc tgcttttgta gaaataacca 2460
tcctttggta tacatggagg atagttccag aacgcctgag tatacaaaaa cccaatgcat 2520
actcaagtcc cacagtgggc cctacagaac ccacgtatgt gataaatcag ccctccatgt 2580
acgcaggttt cgccccctgc caatactgta ttttcaacct gtatggttga aaaaaatcca 2640
tatataagtg cagccatgca gttcaaaccc atattgttca agggtcaact gtatagttta 2700
ttgaacagcc acacccattc ctttacacat gatctatggc agagttgaat agttgcaaca 2760
gacactatgt ggcctgcaaa atcggaaatt tttactgtct ggccttttac agaaaagttt 2820
gccagcccct gatctagacc agcagctcat ctgatagagg cagaggtggc cttaaagatg 2880
tggccttctt catttctgt tggtttggtt tcgtttctat gagagatttc ctctgatagc 2940
tctgctttcc ccagcactta ctctctgagc ttttaaatgt tctctctggg agcttcatat 3000
aagctcggtg acatttgagc cacagttttt agatcagcac ctggaataca tgacacattc 3060
ttactgaggt catccagcac tgccatggtg gctgcccagt cttctggcca gtgtgccagg 3120
cacatgtccc tgtcacacag gttccaagaa acacatacgc agccatgcat agaccaacag 3180
atttaatatt atattgcagt tttcagcgat gcagaatgca gctgcaattg tgttttaagg 3240
agaagccaaa tggggatggt tgtccctgca acatggtgcc actcctgggc catgtgcagc 3300
ctcagtggac actcttccat agcgctgagg ccctggcccc gcctccagtt accctgtact 3360
gcccactgcc ttacagttca gtgcgcaggc cttcaccttt tcatcaccag cctctctgct 3420
cagtgctctg gagttcttga ccttgtcctt tatcatgaga tttgctgaaa tcactaatga 3480
aaataactcc caaaagcaac aaacaaaaat attagtttaa ctggcactgt ggtatattaa 3540
aaggcacaag ggcattgtgg cttaacactt ttgctggatc ccaagagacg cacatgatgt 3600
taaaaagaga tctggcagca gtactaatac tacatttcag tgtaatcatc ttggggtggt 3660
ttggccagga tttcccaatt ccttgatatc tggagtttct tcaccattgt ccggcatcct 3720
gcggaggctt aatatacagg cgtaaggtca gcagcaattt gtctaataag tgatgagatc 3780
agtagctgaa gtctctaagc tgggccatta ctaaatacca tagccatgtt gatctggaaa 3840
tttatccctc tagtgtctta cctcacataa gccatttgcc cactgtgcaa tatagaaagg 3900
tgttttcaaa agtatttggc cgtagatttt cacatccatc ataaggttgg cattcaataa 3960
ggaaaaagtt ctaactccag tattaaattg tacataaatc ccaaatgttc ttaaagaaca 4020
ctcagggaca tgtttgttgc ctgggattgg taatgaaagg ttggtttttg aaacttgaaa 4080
tttcaccatt ggtttttttc ctatcatttc tgcatatcca gcaaaaggaa tctcatgttg 4140
actcctggca gagttcagtg gcttcagtct gtctatctgt tctgagggga aaattgtgtt 4200
ctggatccag taatcaattt ggcaacttta atcgaggttt tcaaaattcc aaggagggtt 4260

aataaagaat gataaccagt tttatttgct aatagctaag acaaatttaa aataaagtgt 4320
tttataatac ttc 4333

<210> 7
<211> 529
<212> PRT
<213> Homo sapiens

<400> 7
Leu Lys Gly Ile Ser Ala Ser Leu Leu Gln Pro Asn Gln Met Pro Ser
1 5 10 15

Ser Thr Ser Pro Asp Gln Gly Asp Asp Leu Glu Asn Cys Ile Leu Arg
20 25 30

Phe Ser Asp Leu Asp Leu Lys Asp Met Ser Leu Ile Asn Pro Ser Ser
35 40 45

Ser Leu Lys Ala Glu Leu Asp Gly Ser Thr Lys Lys Lys Tyr Ser Phe
50 55 60

Ala Lys Lys Lys Ala Phe Ala Leu Phe Val Lys Thr Lys Glu Val Pro
65 70 75 80

Thr Lys Arg Ser Phe Glu Cys Lys Glu Lys Leu Trp Lys Cys Cys Arg
85 90 95

Gln Leu Phe Thr Asp Gln Thr Ser Ile His Arg His Val Ala Thr Gln
100 105 110

His Ala Asp Glu Ile Tyr His Gln Thr Ala Ser Ile Leu Lys Gln Leu
115 120 125

Ala Val Thr Leu Ser Thr Ser Lys Ser Leu Ser Ser Ala Asp Glu Lys
130 135 140

Asn Pro Leu Lys Glu Cys Leu Pro His Ser His Asp Val Ser Ala Trp
145 150 155 160

Leu Pro Asp Ile Ser Cys Phe Asn Pro Asp Glu Leu Ile Ser Gly Gln
165 170 175

Gly Ser Glu Glu Gly Glu Val Leu Leu Tyr Tyr Cys Tyr His Asp Leu
180 185 190

Glu Asp Pro Gln Trp Ile Cys Ala Trp Gln Thr Ala Leu Cys Gln His
195 200 205

Leu His Leu Thr Gly Lys Ile Arg Ile Ala Ala Glu Gly Ile Asn Gly
210 215 220

Thr Val Gly Gly Ser Lys Leu Ala Thr Arg Leu Tyr Val Glu Val Met
225 230 235 240

Leu Ser Phe Pro Leu Phe Lys Asp Asp Leu Cys Lys Asp Asp Phe Lys
245 250 255

Thr Ser Lys Gly Gly Ala His Cys Phe Pro Glu Leu Arg Val Gly Val
260 265 270

Phe Glu Glu Ile Val Pro Met Gly Ile Ser Pro Lys Lys Ile Ser Tyr
275 280 285

Lys Lys Pro Gly Ile His Leu Ser Pro Gly Glu Phe His Lys Glu Val
290 295 300

Glu Lys Phe Leu Ser Gln Ala Asn Gln Glu Gln Ser Asp Thr Ile Leu
305 310 315 320

Leu Asp Cys Arg Asn Phe Tyr Glu Ser Lys Ile Gly Arg Phe Gln Gly
325 330 335

Cys Leu Ala Pro Asp Ile Arg Lys Phe Ser Tyr Phe Pro Ser Tyr Val
340 345 350

Asp Lys Asn Leu Glu Leu Phe Arg Glu Lys Arg Val Leu Met Tyr Cys
355 360 365

Thr Gly Gly Ile Arg Cys Glu Arg Gly Ser Ala Tyr Leu Lys Ala Lys
370 375 380

Gly Val Cys Lys Glu Val Phe Gln Leu Lys Gly Gly Ile His Lys Tyr
385 390 395 400

Leu Glu Glu Phe Pro Asp Gly Phe Tyr Lys Gly Lys Leu Phe Val Phe
405 410 415

Asp Glu Arg Tyr Ala Leu Ser Tyr Asn Ser Asp Val Val Ser Glu Cys
420 425 430

Ser Tyr Cys Gly Ala Arg Trp Asp Gln Tyr Lys Leu Cys Ser Thr Pro
435 440 445

Gln Cys Arg Gln Leu Val Leu Thr Cys Pro Ala Cys Gln Gly Gln Gly
450 455 460

Phe Thr Ala Cys Cys Val Thr Cys Gln Asp Lys Gly Ser Arg Lys Val
465 470 475 480

Ala Gly Pro Met Gln Asp Ser Phe Lys Glu Glu Cys Glu Cys Thr Ala
485 490 495

Arg Arg Pro Arg Ile Pro Arg Glu Leu Leu Gln His Val Arg Gln Pro
500 505 510

Val Ser Pro Glu Pro Gly Pro Asp Ala Asp Glu Asp Gly Pro Val Leu
515 520 525

Met

<210> 8
<211> 4334
<212> DNA
<213> Homo sapiens

<400> 8
atgtccttcc ggggccacgc agaaagtgcc gccgctttgg ccactcagag cccccgggcc 60
gcggtcgtcg tacgcctgaa ggcgggtcgt gccggcggcc gctctagtct ccgcctccgc 120
tcaggccggt cctccggggc ttctcaatgg tttcccggtg gcctctcaat ggttttcccg 180
gcggcccttg cgccgacgcc aggagacttc cggagcttgg tgacgtcacg agcgagcttt 240

```
tctacccaaa tac[redacted]ggcgg gggaataggc tcgagggcgg tgagca[redacted] caattgctag   300
gcggagacag tgcagggaag agagacctta gaaaggatca ggactggcgg gaggtattta   360
actgaaagga atatctgctt cactgttgca accaaaccag atgccttctt ccacttcacc   420
agaccaagga gatgacctgg agaactgcat tttaagattt tctgacctgg atttaaaaga   480
tatgagtctt attaatccca gcagcagtct taaagcagaa ttagatggca gtacaaaaaa   540
gaaatactcg tttgcaaaga aaaaggcctt tgccctttt gtcaaaacca aagaagttcc   600
aacaaaaagg agttttgaat gtaaagaaaa attgtggaaa tgctgtcggc agctattcac   660
agaccaaacc agcatccata gacatgtggc aacacaacat gctgatgaaa tttatcacca   720
gacagcttct attttaaagc aactggctgt gacattgagc acctcaaaga gtctttcgtc   780
tgcagatgaa aagaaccctt taaaagagtg ccttccacat agccatgacg tgtctgcttg   840
gctccctgat ataagctgct ttaaccctga tgagctgata agtggccagg gcagtgaaga   900
aggggaggtg ctcctttatt actgctacca tgacctggag gatccccaat ggatctgtgc   960
ctggcagaca gctctgtgtc agcacctgca cctcacaggc aagattcgaa ttgctgcaga  1020
aggaatcaat gggacagttg gtggaagcaa attggctacc agactttatg tggaagtcat  1080
gctttccttc ccattgttta aggatgacct gtgtaaagat gattttaaga ccagcaaagg  1140
aggagctcac tgttttccag aattgcgtgt tggtgtattt gaagaaatcg tgcccatggg  1200
gatcagcccc aaaaagatct cctacaagaa gcctggaatc catttatccc caggtgaatt  1260
tcataaagaa gtagaaaagt ttttatctca ggcaaatcaa gaacaaagtg atactatcct  1320
tcttgattgc agaaacttct atgaaagcaa aataggacga ttccaaggct gcttagcccc  1380
agacatcagg aaattcagtt acttccctag ctacgttgac aaaaatctag aacttttcag  1440
agagaagaga gtgctgatgt actgtaccgg gggcatccgc tgtgagcggg gttcagccta  1500
cctcaaagcc aagggagtgt gcaaggaggt gttccagctc aagggtggca tccacaagta  1560
cctggaagag tttcctgatg gcttttacaa agggaagttg tttgtttttg atgaacgcta  1620
tgctctgtcc tacaacagtg atgtggtgtc agagtgttca tactgtggag cccgctggga  1680
ccagtataaa ctctgctcta ctcccccagt gccgccagct cgttttgacc tgccctgcct  1740
gtcaaggaca aggattcaca gcctgttgtg tcacatgtca agacaagggg agcaggaaag  1800
tttcaggccc tatgcaagac agctttaaag aggaatgcga gtgcacagcc cgacggccac  1860
gcatacctag ggaactcttg cagcatgtgc gacagcctgt gagcccagag ccagggcctg  1920
atgctgatga ggatgggcca gtgcttatgt gagcagcacc tttggcattt tcccaggccc  1980
tcggtaaaag taggtttggg gtgactatac agagaaagca tggcaagact gcagaaacag  2040
agaaatcggg aacttcagtt ctggccgctg ccaccgtggc agccgtctac acttcacagc  2100
gggaggggag gagtcacgtt gtctaccact tacctgagac attctgattt ggatgatgct  2160
agagcacaga aaataggtga gctgcatggg atcccaaagc tgctgaggga tagagcctga  2220
gcctggtggc cacagcatat gccctttctg ttccatgcag ctggggctgt tagtagtcat  2280
tgcccttgtc agcagacctt ctaccctggt ggcaaacaca tgaaagctgt ggccctggga  2340
gtggcctcct aaaacaagcc acttaggtca tctgccatct acccttaacc tctgtctctc  2400
gcctgagggg aatctgcaag ctgtgcattg ggcttacctc ctgcttttgt agaaataacc  2460
atcctttggt atacatggag gatagttcca gaacgcctga gtatacaaaa acccaatgca  2520
tactcaagtc ccacagtggg ccctacagaa cccacgtatg tgataaatca gccctccatg  2580
tacgcaggtt tcgccccctg ccaatactgt attttcaacc tgtatggttg aaaaaaatcc  2640
atatataagt gcagccatgc agttcaaacc catattgttc aagggtcaac tgtatagttt  2700
attgaacagc cacacccatt cctttacaca tgatctatgg cagagttgaa tagttgcaac  2760
agacactatg tggcctgcaa aatcggaaat ttttactgtc tggcctttta cagaaaagtt  2820
tgccagcccc tgatctagac cagcagctca tctgatagag gcagaggtgg ccttaaagat  2880
gtggccttct tcatttctg ttggtttggt ttcgtttcta tgagagattt cctctgatag  2940
ctctgctttc cccagcactt actctctgag cttttaaatg ttctctctgg gagcttcata  3000
taagctcggt gacatttgag ccacagtttt tagatcagca cctggaatac atgacacatt  3060
cttactgagg tcatccagca ctgccatggt ggctgcccag tcttctggcc agtgtgccag  3120
gcacatgtcc ctgtcacaca ggttccaaga aacacatacg cagccatgca tagaccaaca  3180
gatttaatat tatattgcag ttttcagcga tgcagaatgc agctgcaatt gtgttttaag  3240
gagaagccaa atggggatgg ttgtccctgc aacatggtgc cactcctggg ccatgtgcag  3300
cctcagtgga cactcttcca tagcgctgag gccctggccc cgcctccagt taccctgtac  3360
tgcccactgc cttacagttc agtgcgcagg ccttcacctt ttcatcacca gcctctctgc  3420
tcagtgctct ggagttcttg accttgtcct ttatcatgag atttgctgaa atcactaatg  3480
aaaataactc ccaaaagcaa caaacaaaaa tattagttta actggcactg tggtatatta  3540
aaaggcacaa gggcattgtg gcttaacact tttgctggat cccaagagac gcacatgatg  3600
ttaaaaagag atctggcagc agtactaata ctacatttca gtgtaatcat cttggggtgg  3660
tttggccagg atttcccaat tccttgatat ctggagtttc ttcaccattg tccggcatcc  3720
tgcggaggct taatatacag gcgtaaggtc agcagcaatt tgtctaataa gtgatgagat  3780
cagtagctga agtctctaag ctgggccatt actaaatacc atagccatgt tgatctggaa  3840
atttatccct ctagtgtctt acctcacata agccatttgc ccactgtgca atatagaaag  3900
```

gtgttttcaa aag[redacted]ttgg ccgtagattt tcacatccat cataaggttg gcattcaata 3960
aggaaaaagt tctaactcca gtattaaatt gtacataaat cccaaatgtt cttaaagaac 4020
actcagggac atgtttgttg cctgggattg gtaatgaaag gttggttttt gaaacttgaa 4080
atttcaccat tggttttttt cctatcattt ctgcatatcc agcaaaagga atctcatgtt 4140
gactcctggc agagttcagt ggcttcagtc tgtctatctg ttctgagggg aaaattgtgt 4200
tctggatcca gtaatcaatt tggcaacttt aatcgaggtt ttcaaaattc caaggagggt 4260
taataaagaa tgataatcag ttttatttgc taatagctaa gacaaatttg taataaagtg 4320
ttttataata cttc 4334

<210> 9
<211> 488
<212> PRT
<213> Homo sapiens

<400> 9
Leu Lys Gly Ile Ser Ala Ser Leu Leu Gln Pro Asn Gln Met Pro Ser
1 5 10 15

Ser Thr Ser Pro Asp Gln Gly Asp Asp Leu Glu Asn Cys Ile Leu Arg
20 25 30

Phe Ser Asp Leu Asp Leu Lys Asp Met Ser Leu Ile Asn Pro Ser Ser
35 40 45

Ser Leu Lys Ala Glu Leu Asp Gly Ser Thr Lys Lys Lys Tyr Ser Phe
50 55 60

Ala Lys Lys Lys Ala Phe Ala Leu Phe Val Lys Thr Lys Glu Val Pro
65 70 75 80

Thr Lys Arg Ser Phe Glu Cys Lys Glu Lys Leu Trp Lys Cys Cys Arg
85 90 95

Gln Leu Phe Thr Asp Gln Thr Ser Ile His Arg His Val Ala Thr Gln
100 105 110

His Ala Asp Glu Ile Tyr His Gln Thr Ala Ser Ile Leu Lys Gln Leu
115 120 125

Ala Val Thr Leu Ser Thr Ser Lys Ser Leu Ser Ser Ala Asp Glu Lys
130 135 140

Asn Pro Leu Lys Glu Cys Leu Pro His Ser His Asp Val Ser Ala Trp
145 150 155 160

Leu Pro Asp Ile Ser Cys Phe Asn Pro Asp Glu Leu Ile Ser Gly Gln
165 170 175

Gly Ser Glu Glu Gly Glu Val Leu Leu Tyr Tyr Cys Tyr His Asp Leu
180 185 190

Glu Asp Pro Gln Trp Ile Cys Ala Trp Gln Thr Ala Leu Cys Gln His
195 200 205

Leu His Leu Thr Gly Lys Ile Arg Ile Ala Ala Glu Gly Ile Asn Gly
210 215 220

Thr Val Gly Gly Ser Lys Leu Ala Thr Arg Leu Tyr Val Glu Val Met
225 230 235 240

Leu Ser Phe P[redacted]Leu Phe Lys Asp Asp Leu Cys Lys As[redacted]sp Phe Lys
245 250 255

Thr Ser Lys Gly Gly Ala His Cys Phe Pro Glu Leu Arg Val Gly Val
260 265 270

Phe Glu Glu Ile Val Pro Met Gly Ile Ser Pro Lys Lys Ile Ser Tyr
275 280 285

Lys Lys Pro Gly Ile His Leu Ser Pro Gly Glu Phe His Lys Glu Val
290 295 300

Glu Lys Phe Leu Ser Gln Ala Asn Gln Glu Gln Ser Asp Thr Ile Leu
305 310 315 320

Leu Asp Cys Arg Asn Phe Tyr Glu Ser Lys Ile Gly Arg Phe Gln Gly
325 330 335

Cys Leu Ala Pro Asp Ile Arg Lys Phe Ser Tyr Phe Pro Ser Tyr Val
340 345 350

Asp Lys Asn Leu Glu Leu Phe Arg Glu Lys Arg Val Leu Met Tyr Cys
355 360 365

Thr Gly Gly Ile Arg Cys Glu Arg Gly Ser Ala Tyr Leu Lys Ala Lys
370 375 380

Gly Val Cys Lys Glu Val Phe Gln Leu Lys Gly Gly Ile His Lys Tyr
385 390 395 400

Leu Glu Glu Phe Pro Asp Gly Phe Tyr Lys Gly Lys Leu Phe Val Phe
405 410 415

Asp Glu Arg Tyr Ala Leu Ser Tyr Asn Ser Asp Val Val Ser Glu Cys
420 425 430

Ser Tyr Cys Gly Ala Arg Trp Asp Gln Tyr Lys Leu Cys Ser Thr Pro
435 440 445

Pro Val Pro Pro Ala Arg Phe Asp Leu Pro Cys Leu Ser Arg Thr Arg
450 455 460

Ile His Ser Leu Leu Cys His Met Ser Arg Gln Gly Glu Gln Glu Ser
465 470 475 480

Phe Arg Pro Tyr Ala Arg Gln Leu
485

<210> 10
<211> 21
<212> DNA
<213> Artificial

<220>
<223> expression primer

<400> 10
ctgtgtcagc acctgcacct c 21

<210> 11
<211> 20
<212> DNA
<213> Artificial

<220>
<223> expression primer

<400> 11
atccccatgg gcacgatttc 20

<210> 12
<211> 22
<212> DNA
<213> Artificial

<220>
<223> expression primer

<400> 12
tttccagaat tgcgtgttgg tg 22

<210> 13
<211> 20
<212> DNA
<213> Artificial

<220>
<223> expression primer

<400> 13
tggatgccac ccttgagctg 20

<210> 14
<211> 4888
<212> DNA
<213> Mus musculus

<400> 14
aacccaatta aacttggggg agggagggct ttttttctca acgatgtcag catttatctg 60
tgtgtatcac ctctgagaca ttagcagttg acaaaataaa catcatcatg aactactggt 120
gagcacaata tctcatttag ccccccgtag aaaccaaatg gcttagtatt atccacagtc 180
cacgcttagg tttaagaagt catatgaccc tgacaagatc acatagtaaa taatcgagat 240
gtcgagcttc gaacctaaac ccttcggttc taaagttttt tcttgcctta acccctttgga 300
cctaacacta acagctccaa gtgtgttctt gaaaggcaat aatggtgcaa caaaacaaaa 360
tttatggcaa cacaacgacg tccttgccgg cgaccatcaa aactgacatg agaattaata 420
atctgctcgg gttttcctgc actttcttat aagggaaata tgcttaaagc caatccaaca 480
aaacaccccg tgcctgttta ttagtgacct ctggggtgtt ggtattcttt ttttgaagcc 540
taaaagaggg gaatcagcgg gtggcaggtg gaggagaagg gctcagattg caatacatcc 600
gacctctgtc atagctttgt tcttttcctg tcttcccagg gtcctttcca agtacaccct 660
ggagggctaa aaggaaaaca ctggtccact agatatatct cctccactaa gtgatcctaa 720
gaggcagcag cagacagaca cacgtgcggt ttccatgatc agtagtgagg gattaagtcc 780
tcagtgtcgc gggaccccag gtcggatttg ctcacgggat tcaccaggcc ttcctttccc 840
atgactgaag gagatggggt agcagggagg ccgtcgcttt cccgaagcca ctatccacaa 900
ccgagccagc agcctccctt tggggctccc cgagccctcc cggggagaaa ggattcctat 960
ccgttcccgg gagactgaag gcgagcccgg cctcggttct cacccggacg cagtggccag 1020
gccgagagca cttaccatcg ttctcgtagc tgtgccgcct ccgcgacatg ctgcttgctg 1080
gtgcgcagcc gatctgagtt gctagccgcc tgcaacccca gaaccgtcgg gctgccggcc 1140

ggccggtgcg cg[redacted]agcgg cgaggcgcgc ggactgagcg gcgcgc[redacted]c gcactcgctt 1200
acgacgcctg ccggaagcgc gtgcagaggg gcaccggatg ggctccgcag tggaaggccg 1260
gcgtctgcaa gtccaattgc ggccgcctcc tccggctttc caccccgggc ccttctggac 1320
agcgagctaa attatccggg ccgcagcgat cccgcctcgg ctttatacag gcccggcggg 1380
attttcccgg cggcccctgc gccgacgact taggagactt ccggagaaag gtgacggctc 1440
agggactctc tgcgcacgcg cgcgccgggg ggcgtgggct cgagggcggg gagcagcgcc 1500
aagttggtga agggagaccc aggaaaggcc tagggtttgc gggtaggtgt tctctctttc 1560
tctctcgcca tcccgctgta gacgcctcgg gagacctgag cggggaggag ggcaagccac 1620
tttcagcggt agggaatcca ggcccgtccg gctgaatctg ctcactgtag tcctactgcc 1680
tacttctccc agctgacgga cgccccagct actgatcggt gtggttccaa ttgttctatt 1740
cctctgcacg tgtacagtat gtggacgtcc tctgccctgt gatggacacc aaactgtctt 1800
gtccatggtc tagcccaaac ttttccttca gtcttacctg ttacagtaaa tggccgtttc 1860
ggtagtttgc ttaaatccaa gaatttaggg aggagtcatt agctccttca cttcctgtgt 1920
tacagtttat ccactaacac tgcctgctat actctgagag caagtctagt caaacttatt 1980
tactatgctt aaaactcttc acagacttaa agcaatattt tatgtatttg taagtctgta 2040
catcatcctg gaactggagt ttcagactct ctgtgagcaa ggaggagagt gcgcgggagc 2100
ggggtggggg tgctgaatcc cgggtcttct gctagagcag taagcactct taaccagtga 2160
gccatcttgc catctttccg gctctacttc tctttgttta ttttaaagat tatcgttttt 2220
ctattctttt catttattta ttcgtacatg caagcaatcc atctttccag gaatttccac 2280
tgcgaattcc aaactatggc ttcttgggtt ctgttcccgt cattgtgtgt aaaccaccat 2340
gttcttctgc taaagtccag cttgttcaca ttttaggaat caagacttat tttctttatc 2400
tgttggtctc tgctctgaaa tttacatgac taactgtcag ggacctacag agttatttct 2460
ctgactcctc aactcttgct ttatttcatt cagcaagtac tatattgtat taattaacat 2520
ttgtcaccgt ctgtacatgg gacaacgaga cggtgttcag gatactgaag ccctaagtga 2580
ccctctcagg aatgatcaca ggtactgtag tctgacggac ttaatagcaa agggcgagtg 2640
gtacaaatta tcactcatcc tttgctattg aggacagatt atccagagta agactttgga 2700
agagtcctcc aagaaaggat tttaatatat tctcatctgg ctccagttgc ctgaaatgag 2760
ccaattcgag ccagatgaga atatattaaa aatgggttta ttgggaatct gctctcaggt 2820
gagttcactg accgctacca ggattgagac cagggaagtc actatgggga tgggggtgga 2880
ggagggaaga acacaggagc aaagaaagaa gggaaggagg cccaacaagg aggccaaaag 2940
gtctggattg tatagagagg agcctctggg ggaaggttca gggttgggac agggtatacc 3000
agatagggac tatgggatac tgggagaacc tgaagaccat gtctgctttg atatgtaaaa 3060
tatgcacctt ggtccaaggt tagaaaccaa accttaacca ttactaaatc agataagaag 3120
tgagtgtagt tcaaagtgga ttgctgactt tcaaaggatg gcaaagcatt gggcctggta 3180
tagtggcccc gaattctaga agtagaggca aggggaatct taaatttagg attactagaa 3240
cctcaaggcc agcccctact atatagcaag acccagactc agaaaaatta atttaaaaac 3300
taggggttgg cccaggcagc atacaccagc tgataggagg ccccagacac atatatagca 3360
gaggactgtg tggtctggcc ttagtgaggg aagatgcacc taaccctcaa gaaacttgag 3420
gccccaggga gtgggggcat tctcttgggc aggggacaga gaaggtgtgg gatgaggaac 3480
agagggcaga ctggaaggtg gataaagact agactgtaaa aaagattcaa gaataaaatt 3540
ttcttaattt tgttggaaaa aaaaaatagg ggttggcaag atggctctgt gggcaaaccc 3600
ctgccactca agcctgacaa tacaacgtgt gtttgatccc tggagcctgt gcagatatgg 3660
aactggcagg catctcggca caagtagcca cagaagtaaa tcaaatgtaa caaagaagac 3720
tgcattggtt atagtgtatc cattttggta agatttggcc acagtagcaa tttggaaatt 3780
ttggctctga gatttgagaa caaaagaaac aaagctacag aaacactgga gttggttgat 3840
ttgggaagtc catagtgctt ttctgctctc atgtgctgcc ctgtgttttg ttttcccctg 3900
gcttacttcc tagttcatgt ttctttgaag atgagtgcta ttgcctcaga agggctcttc 3960
ctgtgtatga ggagcccttc cacttgcctc cttactccac tcattttcct ttgaaatatt 4020
taaaaaaaaa aaagaaaaga aaagaaaagt caactcaaat caatgtatcg gcttaccttt 4080
cctagtagat tgtaaactta agaggataga ctttttccta ttgcatttac tgctctattc 4140
ctggtgtgta gaagagagcc tggtagagca tagttgtcac ataaatacac tgtggagtga 4200
gcgaggggac gaacggagag cttggggtag gaggctactg gtcatttcat gagtccatta 4260
gcttcatgtc tgcttcaaga agtaaccaaa gagctgattg ctatattctc ttttgcttac 4320
gtctcagaag atacttgtct gcaaggaacc tgtcccactg ataaagccag atgccttctt 4380
ccacttcacc agacgaagag gatggcctgg agacctgtgt tttaaaggtt tttgatctgg 4440
atttaaaaga atcaaatctt gttaatccca gcaacagtct caaagcagag ttagatggca 4500
gcacaaagaa aaaatactcg tttgcaaaga aaaaggcctt tgccctttg gtcaaaacca 4560
aacaagttcc agcaccctct tatgaattta agggaaaacg gtggcgatgt tgtcagcagc 4620
tgtttgcaga ccagatcagc atccacagac atgtggccac acagcatgct gaagacgtgt 4680
accagcagac tgcgtctctt ctgaagcagc tgactgcagc attgagtgcc tcacagagcc 4740
ttacgcccac agacaaaagg agctccccca aagactgtct cactcctagt caggaggtgt 4800

ctgcttggct tcc[redacted]gtg agccatgtta gcccccagga gctgaggagt ggccagggtg 4860
acgaggaagg agaggtgctc ctgtatta 4888

<210> 15
<211> 7551
<212> DNA
<213> Homo sapiens

<400> 15
gagagacaaa cgttagtaga taaaatttac tcaatttaaa atgtctgtta tgggtttttt 60
tcttacattg acagcatctg ctaacgttta ccgcctctta agcgttggca aatgataaaa 120
catcattgtt gctgatgtgt attacctcat ttaaccttca caaaaaccag aagatataaa 180
tatcattact tctgtattac agactaaagt ttaaggagat ttcataacct ggacaagatc 240
accaagtaaa tggtagggtc tggctttgaa cctaaaccct ctggttccaa agtctcatct 300
cttaaccact actatacact ctcttcaaaa aacaatacac taaaatgtta acaatagaat 360
tactctaagt tttacttttt tggtgtgtgg atttgtgttt ttgaccgatt gtaaggctta 420
tgtgtaatat aagcatcttt tttttgttgt tgttaaaact agaattattc catctttatg 480
aaggcaacac tgatctccac gagcacaatc atccaaacgg agatgggaac ccctgttcta 540
actctctcag aggcctgaag ttagttttct agcactctct ctcacacaag ggaaatgcgc 600
ttacggccag ttcaaaatct gcagtaacct gtgtctgttt attagggacc tcttgggcgc 660
aggtaagaca atattctttc cctgcagctt gaaagaagtg aaatcaacgg ataaccacta 720
ccggaacaga gttcagactg caactcatga gatctctcag tcctttcctt cagtgaattc 780
ttccgagccc ttttcaggtg cagctttgaa ggttaaaacg agacactagg accagtgggt 840
ccctaagagg cggcagcaca gacacaggtg tgcttccctg ctccgtggtt ttacggacga 900
aattctgaag catccttaga gaccaagggt ctacccatgg ggttcaccag gccttctttc 960
ccgaggatcg gcccttgact caaagaactg gggcggggga gaagtcgctt tctagaagcc 1020
gcccgaccac accgccacca tattctcctc tttccctccg cggctccccg cacgttcccg 1080
gggagaagag cttccagtct cttccgtgga cacccaaaga cgccgagcct cggctcccaa 1140
ccgggagcgg cctccgtggc cgggccgcag gcactcaccg tcgttctcgt cgctgtgccg 1200
ccgccgcgac atgctgccct ccggtgcgcc gcggaaccga gaggccaggc ggtaagcgct 1260
gcaggaactg tctggccgct ggccgacgca aggacagctg caaggcgcgc ggatgggccg 1320
gcacgcaggc gcactagctc gccacggccc cggaagcgca ggagaggccg ccgggtgggg 1380
ctaggcgctg cgacagccgg cgtgaggaag ctcagtgggc tacgaacgtc tggcacacat 1440
gcaaccgccc cctcgggctg cctccgcctg ccggctactt ctttctcccg ccttccgctc 1500
tatgtccttc cggggccacg cagaaagtgc cgccgctttg gccactcaga gcccccgggc 1560
cgcggtcgtc gtacgcctga aggcgggtcg tgccggcggc cgctctagtc tccgcctccg 1620
ctcaggccgg tcctccgggg cttctcaatg gtttcccggt ggcctctcaa tggttttccc 1680
ggcggccctt gcgccgacgc caggagactt ccggagcttg gtgacgtcac gagcgagctt 1740
ttctacccaa atacgcggcg ggggaatagg ctcgagggcg gtgagcagtg acaattgcta 1800
ggcggagaca gtgcagggaa gagagacctt agaaaggatc aggactggcg ggtatgtgct 1860
catctactcc cactttccgg cttttgccgc cttgggaaaa gtgggaggag aggttgggcc 1920
aagctcggca tgcggggtgg ggcctgggcg ggaggcggtg ccgcacgtgc cgccccttgg 1980
tatggaaagg ccggccctga cgcgagcgtg cggctccggg cttgccggct ggccgtcatt 2040
ttccttaagt tgtttgctta ggaaggaaca aatgattgtt ttagtaatct gttttcaagg 2100
gttattgggt acctatatgc ctgtgtgtga gcgtcccctc ccctgatgtc tttatgagca 2160
ccatactggc ttatctgctt tgtccatggc catctgaaat ttcactctcc cggcctaatt 2220
ctccttcggt cttacccatt tcagtaaatg acgtcatccg tcggccgctt aatccaaaaa 2280
tttagagtct ttcttagctc ttttccctct ttttacataa caattatctg taaaccctgc 2340
ctgctctgct ttcaacacaa atccagaatc agatcgtgtt actgctttgt ttaaaagtct 2400
cctatgacgg ccgggcgggg tagctaacgc ttgtaatcct aacacttatg ggaggccgac 2460
gcgggcggat ccgcttgagc ctaggagttc aagacccacc tgggcaacaa ggcgaacacc 2520
gtctttacag gaaaaaaaaa aaaattagct gagcctagtg tttggcgcct gtagtcccag 2580
ttactcgagg ggagggttgg gggccgaggt gggaggatcc cttgagccca ggaggtcgag 2640
aatgcagtga gcggtgatcg cacactgcac tccagcctgg gcaacaaagt gagaccgtgt 2700
caaaaaaaaa aagtctccta cgactttcca ttgcacactc catgccatgg cctgtggggt 2760
tctacttctc ctcactgtac tccaaccact ctttctgttc ttacaataag tcaagcttgt 2820
tcttatttta ggaacttata ctatttcctt tacctgaagg tctctggttc taaatctgcg 2880
tggctggctc ttggtcactt aggtctcagc tcaaatgtca ggtccttagt gaggccttct 2940
ttggttaccc aatcactgtt ttatttcatt caacgaatat tatatttatt agcatttagc 3000
acctcctagg tgccaggcag agttctgaat gctgaggatg tagaggtgaa caaaggaaaa 3060

```
cccctgctct ct[redacted]cgctg atatattcta acgacccatt cagga[redacted]c cacaaatagt   3120
gccatcatag gtagtgagca acgggagagt ggtagatgag atcagaggagg taggggggaag   3180
gtgctctgtg aggactttgg acattatcca agtgaattga gaagtcattg gagtatgatt   3240
ttttttttt gagacagggt ctcgctctgt tgcccaggct ggagtgcagt ggcacggtct   3300
tggctcgttg taacctctgc cttttaggtt caggcaattc ttgtgcctta gccacccaag   3360
tagcggcaat tacagacatg cgccaccaca ccaggctaat ttgtatattt ttagtagaga   3420
tggggtttcg ccatgttggc caggctggtc ttgaactcct ggcctcccaa agtgctgaga   3480
ttacaggcat gagcctctgt gcagggccag tctgacttat ttttttaata gttacttctc   3540
atgctcccaa aggaggaaga tagttcagct tttgcagtaa ttctagctgc aaatgatagt   3600
agcttgaact agggtggtag taaagatggt ggtgcaaagt gatctgattg gattgtactt   3660
caaaagcaaa gctggccagg catggtggct cacccctgta atcctagcac tttgggagtc   3720
cgaggtgggt ggatcacaag gtctggagat tgagaccatc ctggccaaca tggtgaaacc   3780
ctgtctctac taaaaaaaaa aatacaaaaa ttagctgcgt gtggtggcgt gcagctgtac   3840
tcccagctac tcgggaggct gaggcaggag aatcgcttga acctgggagg cggaggttgc   3900
agtgagccga gatcacccca ctgcactcca gcctggtgac agaacgaaaa ccgtctcaaa   3960
aaaaagtatt ttctgaacaa gtgaatgata gtgtgagaaa ataaggagtc aggaatcatc   4020
ccaaagtttt ttttgtctga ataacaggaa gaataaaatt gtcttttact cagataagga   4080
agactgggga agacgaaagt ttggttttga acagttgtat ttgaagctga agctagctgg   4140
cttcacacac agtgtaaaat tcacacacag tggggtatat gaatttgtag tttgggagaa   4200
aggcccaggt tgaagatata aatatggagt catttattat atagtattct aaacagaaaa   4260
tttcctttat gtttcccttt tttcttttct catgtaattt tcaaatattt acagaagtag   4320
ggaaaattta tgatgaaacc ccatgtactc atcatcaact tcagtaattc ttaattcatg   4380
ggtaatctgg tttcattgat attacctcct ctctctcaat tattttgagg caaatatagt   4440
tttaagtcct tggaagtaga tgagatcacc tgtagggtaa gtatagatag agtggttcaa   4500
catttagagg ttggggtaat aaggaagatc cagcaaagga aactgagaag caggggccgt   4560
tatccctagc tggtattttc atgatttgtt tgtttgtttg tttgtttgtt tattacttac   4620
ttaatgtctc tccttaccag aatataagct ccaagagagc aggtacttca actgtttcat   4680
tcatagctgc attcccagga ctattccatg ggcccggcac atagcagagg tgccacaaat   4740
aataataatt cagagtatgt acataaatca aaataagctg agactgaaaa aagatgcaga   4800
gtcggaggag atcattttgg ggtggtggga aaggcacccc acactcaggt gcagagacct   4860
tgagtcagga aagaacatgg ctcatttgtg gaactgggca tggtccagtg tggctagagt   4920
aggatatctg atggggaatg gggagcatag tcagagatga gcctagagta tagggaaact   4980
aagtttttca ggccttcaaa ggaagcagtg acacagggta gcagacaaga cataagtgaa   5040
aggtgacact tgagttcagt cttcatatac atagtaagca tatggggagg gaaagaagtg   5100
aaaaagtaag ggaaacaaag cagtggaaag cataaatgct gggacagact acatcaggtc   5160
atatagaggg agattggaaa agatgctgga ggtgggaccg gattggatag taattccttt   5220
cgacaatgtt acctcccaaa tgtactgccc ttataggtaa ataagttatc tgtttctaga   5280
acaatgcctg gaacctagaa gaccctccat aaatatttgt tgaatgaatt ttaatttgga   5340
aatgatatat gtactttctc tacctttaca ataaaacagc tttgagcatg gaggaagtgt   5400
gttgaataag attcttctac ttattttact tctgcctgga aggcttacag gtttgtgtag   5460
agttagggat gtacttttt cagtgggact ttatatgaaa gacattttca tgagaaagga   5520
aagagaagct aaagaaggta gtatttcttc acccagtctc ttccacttat ttttcaaata   5580
caagagaatt taggtttaat aaatctgtat taccttgcct tacccttata acaaagctca   5640
ttttgagccc catgaaagca ccacttagct gattcaaatg cattttatgt ggtgcaagtg   5700
acatcttaat ttatgaatcc aaggaacaaa ggtcctttat aagaaagttg gtattgcctg   5760
tgtgtgtctg tgtatgttta gtctccatca gtcctatcca gcattgttaa caagcagtgt   5820
gtgttccccc acacataaac tgctccttgt tgccttcacc tggttgaggt accctgtagg   5880
cagagttgtg aattggcaaa gtttgtcacc agtgtctccc accttccttt ctctcattct   5940
tagtggtgct ttttctgctc cctcttgtgg agtttgggac tactttgcat gttactattt   6000
tcagaacctt cagtgtccac ctctggctgc ctccttcaat attgtcctga agaaactaac   6060
tttgccagtt cactagtaat atgaaataac cagttgttat gttgtggggc aggaatttgg   6120
acagggcaca acagggacac cttgtctctg ctctcagctc agaggccaat ctggaatctg   6180
aagggtctag tggttgatgc tggctgtcag ctggaggcct ctgttcctct ccacctggtt   6240
ctctgtgtgt ggtctctgca tgggttggat tggcttactc acaatacagt ggtttggtct   6300
cctcagcgtg agcatctcag aatgagagct tgcactagcg tatccttttg agaacccaac   6360
ctcagtcata cagcattgct tccatcacgc tttatttttg ggagtagtta acaagcctgc   6420
ctaggcttag ggaaatggac cagaaatact gctgcagcca tttttggaaa atataatctg   6480
tcacagctaa tttttaaata aaattttaat gttaaattta aaatacagca ttcaattaaa   6540
atattaataa tattaccatt cattgaaccc tagtttgtgc cagacactat atattaatat   6600
ttctttccag tctacagtac agtactgtaa tgtagatact acaatgcaca atttccacaa   6660
gagagattta gaactcagaa accttaagta gcttgcccat gttcacatgc ttggcaaact   6720
```

gctgcagcac aaa[redacted]agc ccatactttt tctcagtgtg cctcttag[redacted]cttgtcacct 6780
gtcacctttg ctttacaaaa gaaccaacat ggtatgatgg ttgtagtgtg tccattttga 6840
caagatttgg ccacagtagc gattttgaca tttcttctgg catattacag tggggctgtc 6900
ttgggaacta tttcttctaa ggaaggaaat tttcatagct ctgaaattaa ggggcaaaaa 6960
gaagaaagtg tttataactg aaggtgctgg cacacatctt gattgatttg ggttgtacct 7020
gttgcttttc tggtattatt taccctaatg actgcatggc tcgctcccta attcaggttt 7080
ctgctcaaat gctattgctg cagaaaggct tttcctactc tctgtgttcg aagggcttct 7140
ccatcactgc ctccctaccc tgctgattat tcttcataac ctttacacca tctgacttta 7200
aaaataaaaa atcaatttat ccacttgttt acttcacccc actggaatgt aaactctaag 7260
caggcagaaa ctttttcact ttcgtttact gctatattct cagtgcttag aatagtgctt 7320
gacttcgagt cgttcctaag taaatactat tgatttagta agtaaatgga tggttagctt 7380
gggctacttt gcccatgctg atttgcgtga ttgattagct tcatattttc ttcaagaaat 7440
caaagaatag catggctgct tctttcttat ttatattcta ggaggtattt aactgaaagg 7500
aatatctgct tcactgttgc aaccaaacca gatgccttct tccacttcac c 7551

<210> 16
<211> 488
<212> PRT
<213> Homo sapiens

<400> 16
Leu Lys Gly Ile Ser Ala Ser Leu Leu Gln Pro Asn Gln Met Pro Ser
1 5 10 15

Ser Thr Ser Pro Asp Gln Gly Asp Asp Leu Glu Asn Cys Ile Leu Arg
20 25 30

Phe Ser Asp Leu Asp Leu Lys Asp Met Ser Leu Ile Asn Pro Ser Ser
35 40 45

Ser Leu Lys Ala Glu Leu Asp Gly Ser Thr Lys Lys Lys Tyr Ser Phe
50 55 60

Ala Lys Lys Lys Ala Phe Ala Leu Phe Val Lys Thr Lys Glu Val Pro
65 70 75 80

Thr Lys Arg Ser Phe Glu Cys Lys Glu Lys Leu Trp Lys Cys Cys Arg
85 90 95

Gln Leu Phe Thr Asp Gln Thr Ser Ile His Arg His Val Ala Thr Gln
100 105 110

His Ala Asp Glu Ile Tyr His Gln Thr Ala Ser Ile Leu Lys Gln Leu
115 120 125

Ala Val Thr Leu Ser Thr Ser Lys Ser Leu Ser Ser Ala Asp Glu Lys
130 135 140

Asn Pro Leu Lys Glu Cys Leu Pro His Ser His Asp Val Ser Ala Trp
145 150 155 160

Leu Pro Asp Ile Ser Cys Phe Asn Pro Asp Glu Leu Ile Ser Gly Gln
165 170 175

Gly Ser Glu Glu Gly Glu Val Leu Leu Tyr Tyr Cys Tyr His Asp Leu
180 185 190

Glu Asp Pro Gln Trp Ile Cys Ala Trp Gln Thr Ala Leu Cys Gln His
195 200 205

Leu His Leu Th[redacted]Gly Lys Ile Arg Ile Ala Ala Glu Gl[redacted]e Asn Gly
210 215 220

Thr Val Gly Gly Ser Lys Leu Ala Thr Arg Leu Tyr Val Glu Val Met
225 230 235 240

Leu Ser Phe Pro Leu Phe Lys Asp Asp Leu Cys Lys Asp Asp Phe Lys
245 250 255

Thr Ser Lys Gly Gly Ala His Cys Phe Pro Glu Leu Arg Val Gly Val
260 265 270

Phe Glu Glu Ile Val Pro Met Gly Ile Ser Pro Lys Lys Ile Ser Tyr
275 280 285

Lys Lys Pro Gly Ile His Leu Ser Pro Gly Glu Phe His Lys Glu Val
290 295 300

Glu Lys Phe Leu Ser Gln Ala Asn Gln Glu Gln Ser Asp Thr Ile Leu
305 310 315 320

Leu Asp Cys Arg Asn Phe Tyr Glu Ser Lys Ile Gly Arg Phe Gln Gly
325 330 335

Cys Leu Ala Pro Asp Ile Arg Lys Phe Ser Tyr Phe Pro Ser Tyr Val
340 345 350

Asp Lys Asn Leu Glu Leu Phe Arg Glu Lys Arg Val Leu Met Tyr Cys
355 360 365

Thr Gly Gly Ile Arg Cys Glu Arg Gly Ser Ala Tyr Leu Lys Ala Lys
370 375 380

Gly Val Cys Lys Glu Val Phe Gln Leu Lys Gly Gly Ile His Lys Tyr
385 390 395 400

Leu Glu Glu Phe Pro Asp Gly Phe Tyr Lys Gly Lys Leu Phe Val Phe
405 410 415

Asp Glu Arg Tyr Ala Leu Ser Tyr Asn Ser Asp Val Val Ser Glu Cys
420 425 430

Ser Tyr Cys Gly Ala Arg Trp Asp Gln Tyr Lys Leu Cys Ser Thr Pro
435 440 445

Pro Val Pro Pro Ala Arg Phe Asp Leu Pro Cys Leu Ser Arg Thr Arg
450 455 460

Ile His Ser Leu Leu Cys His Met Ser Arg Gln Gly Glu Gln Glu Ser
465 470 475 480

Phe Arg Pro Tyr Ala Arg Gln Leu
485

<210> 17
<211> 529
<212> PRT
<213> Homo sapiens

<400> 17

Leu Lys Gly Ile Thr Ala Ser Leu Leu Gln Pro Asn Gln Met Pro Ser
1 10 15

Ser Thr Ser Pro Asp Gln Gly Asp Asp Leu Glu Asn Cys Ile Leu Arg
20 25 30

Phe Ser Asp Leu Asp Leu Lys Asp Met Ser Leu Ile Asn Pro Ser Ser
35 40 45

Ser Leu Lys Ala Glu Leu Asp Gly Ser Thr Lys Lys Lys Tyr Ser Phe
50 55 60

Ala Lys Lys Lys Ala Phe Ala Leu Phe Val Lys Thr Lys Glu Val Pro
65 70 75 80

Thr Lys Arg Ser Phe Glu Cys Lys Glu Lys Leu Trp Lys Cys Cys Arg
85 90 95

Gln Leu Phe Thr Asp Gln Thr Ser Ile His Arg His Val Ala Thr Gln
100 105 110

His Ala Asp Glu Ile Tyr His Gln Thr Ala Ser Ile Leu Lys Gln Leu
115 120 125

Ala Val Thr Leu Ser Thr Ser Lys Ser Leu Ser Ser Ala Asp Glu Lys
130 135 140

Asn Pro Leu Lys Glu Cys Leu Pro His Ser His Asp Val Ser Ala Trp
145 150 155 160

Leu Pro Asp Ile Ser Cys Phe Asn Pro Asp Glu Leu Ile Ser Gly Gln
165 170 175

Gly Ser Glu Glu Gly Glu Val Leu Leu Tyr Tyr Cys Tyr His Asp Leu
180 185 190

Glu Asp Pro Gln Trp Ile Cys Ala Trp Gln Thr Ala Leu Cys Gln His
195 200 205

Leu His Leu Thr Gly Lys Ile Arg Ile Ala Ala Glu Gly Ile Asn Gly
210 215 220

Thr Val Gly Gly Ser Lys Leu Ala Thr Arg Leu Tyr Val Glu Val Met
225 230 235 240

Leu Ser Phe Pro Leu Phe Lys Asp Asp Leu Cys Lys Asp Asp Phe Lys
245 250 255

Thr Ser Lys Gly Gly Ala His Cys Phe Pro Glu Leu Arg Val Gly Val
260 265 270

Phe Glu Glu Ile Val Pro Met Gly Ile Ser Pro Lys Lys Ile Ser Tyr
275 280 285

Lys Lys Pro Gly Ile His Leu Ser Pro Gly Glu Phe His Lys Glu Val
290 295 300

Glu Lys Phe Leu Ser Gln Ala Asn Gln Glu Gln Ser Asp Thr Ile Leu
305 310 315 320

Leu Asp Cys Arg Asn Phe Tyr Glu Ser Lys Ile Gly Arg Phe Gln Gly

325 330 335

Cys Leu Ala Pro Asp Ile Arg Lys Phe Ser Tyr Phe Pro Ser Tyr Val
340 345 350

Asp Lys Asn Leu Glu Leu Phe Arg Glu Lys Arg Val Leu Met Tyr Cys
355 360 365

Thr Gly Gly Ile Arg Cys Glu Arg Gly Ser Ala Tyr Leu Lys Ala Lys
370 375 380

Gly Val Cys Lys Glu Val Phe Gln Leu Lys Gly Gly Ile His Lys Tyr
385 390 395 400

Leu Glu Glu Phe Pro Asp Gly Phe Tyr Lys Gly Lys Leu Phe Val Phe
405 410 415

Asp Glu Arg Tyr Ala Leu Ser Tyr Asn Ser Asp Val Val Ser Glu Cys
420 425 430

Ser Tyr Cys Gly Ala Arg Trp Asp Gln Tyr Lys Leu Cys Ser Thr Pro
435 440 445

Gln Cys Arg Gln Leu Val Leu Thr Cys Pro Ala Cys Gln Gly Gln Gly
450 455 460

Phe Thr Ala Cys Cys Val Thr Cys Gln Asp Lys Gly Ser Arg Lys Val
465 470 475 480

Ser Gly Pro Met Gln Asp Ser Phe Lys Glu Glu Cys Glu Cys Thr Ala
485 490 495

Arg Arg Pro Arg Ile Pro Arg Glu Leu Leu Gln His Val Arg Gln Pro
500 505 510

Val Ser Pro Glu Pro Gly Pro Asp Ala Asp Glu Asp Gly Pro Val Leu
515 520 525

Met

<210> 18
<211> 555
<212> DNA
<213> Homo sapiens

<400> 18
aaggcgcgcg gatgggccgg cacgcaggcg cactagctcg ccacggcccc ggaagcgcag 60
gagaggccgc cgggtggggc taggcgctgc gacagccggc gtgaggaagc tcagtgggct 120
acgaacgtct ggcacacatg caaccgcccc ctcgggctgc ctccgcctgc cggctacttc 180
tttctcccgc cttccgctca tgtccttccg gggccacgca gaaagtgccg ccgctttggc 240
cactcagagc ccccgggccg cggtcgtcgt acgcctgaag gcgggtcgtg ccggcggccg 300
ctctagtctc cgcctccgct caggccggtc ctccggggct tctcaatggt ttcccggtgg 360
cctctcaatg gttttcccgg cggcccttgc gccgacgcca ggagacttcc ggagcttggt 420
gacgtcacga gcgagctttt ctacccaaat acgcggcggg ggaataggct cgagggcggt 480
gagcagtgac aattgctagg cggagacagt gcagggaaga gagaccttag aaaggatcag 540
gactggcggg tatgt 555

<210> 19
<211> 237

<212> DNA
<213> Homo sapiens

<400> 19
cttatttata ttctaggagg tatttaactg aaaggaatat ctgcttcact gttgcaacca 60
aaccagatgc cttcttccac ttcaccagac caaggagatg acctggagaa ctgcatttta 120
agattttctg acctggattt aaaagatatg agtcttatta atcccagcag cagtcttaaa 180
gcagaattag atggcagtac aaaaaagaaa tactcgtttg caaagaaaaa ggtagaa 237

<210> 20
<211> 339
<212> DNA
<213> Homo sapiens

<400> 20
tgaaatctaa ttgcaggcct ttgccctttt tgtcaaaacc aaagaagttc caacaaaaag 60
gagttttgaa tgtaaagaaa aattgtggaa atgctgtcgg cagctattca cagaccaaac 120
cagcatccat agacatgtgg caacacaaca tgctgatgaa atttatcacc agacagcttc 180
tattttaaag caactggctg tgacattgag cacctcaaag agtctttcgt ctgcagatga 240
aaagaaccct ttaaaagagt gccttccaca tagccatgac gtgtctgctt ggctccctga 300
tataagctgc tttaaccctg atgagctgat aaggtaaga 339

<210> 21
<211> 143
<212> DNA
<213> Homo sapiens

<400> 21
gattttcatt ttatagtggc cagggcagtg aagaagggga ggtgctcctt tattactgct 60
accatgacct ggaggatccc caatggatct gtgcctggca gacagctctg tgtcagcacc 120
tgcacctcac aggcaaggta aca 143

<210> 22
<211> 148
<212> DNA
<213> Homo sapiens

<400> 22
ccgtcttgtg tctcagattc gaattgctgc agaaggaatc aatgggacag ttggtggaag 60
caaattggct accagacttt atgtggaagt catgctttcc ttcccattgt ttaaggatga 120
cctgtgtaaa gatgatttta aggtaaga 148

<210> 23
<211> 128
<212> DNA
<213> Homo sapiens

<400> 23
gtttctcatt ggctagacca gcaaaggagg agctcactgt tttccagaat tgcgtgttgg 60
tgtatttgaa gaaatcgtgc ccatggggat cagccccaaa aagatctcct acaagaagcc 120
tggtatgc 128

<210> 24
<211> 141
<212> DNA
<213> Homo sapiens

<400> 24
tttggtttgg ttttaggaat ccatttatcc ccaggtgaat ttcataaaga agtagaaaag 60
tttttatctc aggcaaatca agaacaaagt gatactatcc ttcttgattg cagaaacttc 120
tatgaaagca aaatagtaag t 141

<210> 25
<211> 181
<212> DNA
<213> Homo sapiens

<400> 25
tgctcctatg ttacagggac gattccaagg ctgcttagcc ccagacatca ggaaattcag 60
ttacttccct agctacgttg acaaaaatct agaacttttc agagagaaga gagtgctgat 120
gtactgtacc gggggcatcc gctgtgagcg gggttcagcc tacctcaaag ccaaggtgag 180
c 181

<210> 26
<211> 161
<212> DNA
<213> Homo sapiens

<400> 26
gtttttccac acctagggag tgtgcaagga ggtgttccag ctcaagggtg gcatccacaa 60
gtacctggaa gagtttcctg atggctttta caaagggaag ttgtttgttt ttgatgaacg 120
ctatgctctg tcctacaaca gtgatgtggt gtcaggtagg t 161

<210> 27
<211> 2697
<212> DNA
<213> Homo sapiens

<400> 27
tttccttcct ccccagagtg ttcatactgt ggagcccgct gggaccagta taaactctgc 60
tctactcccc agtgccgcca gctcgttttg acctgccctg cctgtcaagg acaaggattc 120
acagcctgtt gtgtcacatg tcaagacaag gggagcagga aagtttcagg ccctatgcaa 180
gacagcttta aagaggaatg cgagtgcaca gcccgacggc cacgcatacc tagggaactc 240
ttgcagcatg tgcgacagcc tgtgagccca gagccagggc ctgatgctga tgaggatggg 300
ccagtgctta tgtgagcagc acctttggca ttttcccagg ccctcggtaa aagtaggttt 360
ggggtgacta tacagagaaa gcatggcaag actgcagaaa cagagaaatc gggaacttca 420
gttctggccg ctgccaccgt ggcagccgtc tacacttcac agcgggaggg gaggagtcac 480
gttgtctacc acttacctga gacattctga tttggatgat gctagagcac agaaaatagg 540
tgagctgcat gggatcccaa agctgctgag ggatagagcc tgagcctggt ggccacagca 600
tatgcccttt ctgttccatg cagctggggc tgttagtagt cattgccctt gtcagcagac 660
cttctaccct ggtggcaaac acatgaaagc tgtggccctg ggagtggcct cctaaaacaa 720
gccacttagg tcatctgcca tctaccctta acctctgtct ctcgcctgag gggaatctgc 780
aagctgtgca ttgggcttac ctcctgcttt tgtagaaata accatccttt ggtatacatg 840
gaggatagtt ccagaacgcc tgagtataca aaaacccaat gcatactcaa gtcccacagt 900
gggccctaca gaacccacgt atgtgataaa tcagccctcc atgtacgcag gtttcgcccc 960
ctgccaatac tgtattttca acctgtatgg ttgaaaaaaa tccatatata agtgcagcca 1020
tgcagttcaa acccatattg ttcaagggtc aactgtatag tttattgaac agccacaccc 1080
attcctttac acatgatcta tggcagagtt gaatagttgc aacagacact atgtggcctg 1140
caaaatcgga aatttttact gtctggcctt ttacagaaaa gtttgccagc ccctgatcta 1200
gaccagcagc tcatctgata gaggcagagg tggccttaaa gatgtggcct tcttcatttt 1260
ctgttggttt ggtttcgttt ctatgagaga tttcctctga tagctctgct ttccccagca 1320
cttactctct gagcttttaa atgttctctc tgggagcttc atataagctc ggtgacattt 1380
gagccacagt ttttagatca gcacctggaa tacatgacac attcttactg aggtcatcca 1440
gcactgccat ggtggctgcc cagtcttctg gccagtgtgc caggcacatg tccctgtcac 1500

acaggttcca agaaaaacat acgcagccat gcatagacca acagatttaa tattatattg 1560
cagttttcag cgatgcagaa tgcagctgca attgtgtttt aaggagaagc caaatgggga 1620
tggttgtccc tgcaacatgg tgccactcct gggccatgtg cagcctcagt ggacactctt 1680
ccatagcgct gaggccctgg ccccgcctcc agttaccctg tactgcccac tgccttacag 1740
ttcagtgcgc aggccttcac cttttcatca ccagcctctc tgctcagtgc tctggagttc 1800
ttgaccttgt cctttatcat gagatttgct gaaatcacta atgaaaataa ctcccaaaag 1860
caacaaacaa aaatattagt ttaactggca ctgtggtata ttaaaaggca caagggcatt 1920
gtggcttaac acttttgctg gatcccaaga gacgcacatg atgttaaaaa gagatctggc 1980
agcagtacta atactacatt tcagtgtaat catcttgggg tggtttggcc aggatttccc 2040
aattccttga tatctggagt ttcttcacca ttgtccggca tcctgcggag gcttaatata 2100
caggcgtaag gtcagcagca atttgtctaa taagtgatga gatcagtagc tgaagtctct 2160
aagctgggcc attactaaat accatagcca tgttgatctg gaaatttatc cctctagtgt 2220
cttacctcac ataagccatt tgcccactgt gcaatataga aaggtgtttt caaaagtatt 2280
tggccgtaga ttttcacatc catcataagg ttggcattca ataaggaaaa agttctaact 2340
ccagtattaa attgtacata aatcccaaat gttcttaaag aacactcagg gacatgtttg 2400
ttgcctggga ttggtaatga aaggttggtt tttgaaactt gaaatttcac cattggtttt 2460
tttcctatca tttctgcata tccagcaaaa ggaatctcat gttgactcct ggcagagttc 2520
agtggcttca gtctgtctat ctgttctgag gggaaaattg tgttctggat ccagtaatca 2580
atttggcaac tttaatcgag gttttcaaaa ttccaaggag ggttaataaa gaatgataat 2640
cagttttatt tgctaatagc taagacaaat ttgtaataaa gtgttttata atacttc 2697

<210> 28
<211> 300
<212> DNA
<213> Homo sapiens

<400> 28
gtgcttttct ctttaggtta cgagacagta caatagaagg agtatgctcg tccccattct 60
ttcactgagt caccatatga ttttggacca gctagtgctc tagacctcag tatcccttct 120
tataaaataa gaatgttaca gctcatgcaa tctgggactc caaatcttgg acatattagc 180
tcacttgaga gaccaccagc ctggtcagca gatcactgtg tttttagtaa atctggaatt 240
gtaagattaa cacttcatac cacatggggg aataaagttg ttgctctcac aggtgggctg 300

<210> 29
<211> 813
<212> DNA
<213> Homo sapiens

<400> 29
gtttttccac acctagggag tgtgcaagga ggtgttccag ctcaagggtg gcatccacaa 60
gtacctggaa gagtttcctg atggctttta caaagggaag ttgtttgttt ttgatgaacg 120
ctatgctctg tcctacaaca gtgatgtggt gtcaggtagg tcagcacagg ctcagagccc 180
aaactgaaat gaagcacatt gtcagttcac tattctagaa aaatgacaca gggaagacag 240
gccagtgctc attactgagc actgaataag cagggaaaat aagtacattg tgccaccatt 300
ttcccagctg tggagctgag agaaccctag cccaggagtc aggaggcctg ggttgggatc 360
ctggcttcac cattgctagc tggacaagcc cattaacatg gggatcatct cacctgccct 420
gcctgcctgt ctacctgcca agagctgtac tactgggcta attcagggct cttaacctgg 480
aattggtaca tagatttcag ggattctgtg aatttggatg gaaaaataat tgtatctttg 540
ttttcaataa cacctcacta aaatgaagca tttcctttag ttatgaatgt aggcaacaaa 600
gtaccagttg tattaatgta cctgtgactt tgtcttcagt aggattcaca atactttcat 660
atcatgttct agttgcctca gatatctcaa aatagtattt atactcatca ctgcttcaaa 720
atgaaaatag ttattaggcc caccactaag agttgatata taatgtgtta ataaatggca 780
cgtcttatta tatattacag attttgaaaa aga 813

<210> 30
<211> 5
<212> DNA
<213> Artificial

<400> 30
ctttga 5

<210> 31
<211> 18
<212> DNA
<213> Artificial

<220>
<223> expression primer

<400> 31
ccaagggagt gtgcaagg 18

<210> 32
<211> 25
<212> DNA
<213> Artificial

<220>
<223> expression primer

<400> 32
cctttgtaaa agccatcagg aaact 25